RESEARCH PAPER

On partial likelihood and the construction of factorisable transformations

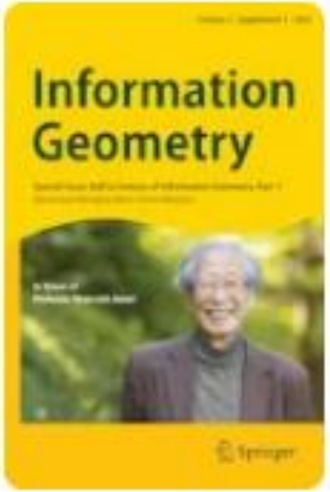
H. S. Battey[1] · D. R. Cox[2] · Su Hyeong Lee[3]

Received: 15 April 2022 / Revised: 11 June 2022 / Accepted: 15 June 2022 /
Published online: 28 June 2022

Abstract

Models whose associated likelihood functions fruitfully factorise are an important minority allowing elimination of nuisance parameters via partial likelihood, an operation that is valuable in both Bayesian and frequentist inferences, particularly when the number of nuisance parameters is not small. After some general discussion of partial likelihood, we focus on marginal likelihood factorisations, which are particularly difficult to ascertain from elementary calculations. We suggest a systematic approach for deducing transformations of the data, if they exist, whose marginal likelihood functions are free of the nuisance parameters. This is based on the solution to an integro-differential equation constructed from aspects of the Laplace transform of the probability density function, for which candidate solutions solve a simpler first-order linear homogeneous differential equation. The approach is generalised to the situation in which such factorisable structure is not exactly present. Examples are used in illustration. Although motivated by inferential problems in statistics, the proposed construction is of independent interest and may find application elsewhere.

Keywords Inferential separation · Marginal likelihood · Matched comparisons · Method of characteristics · Partial differential equations · Nuisance parameters



Volume 7
Special Issue 1
Article 1
Number 93

PERSPECTIVE

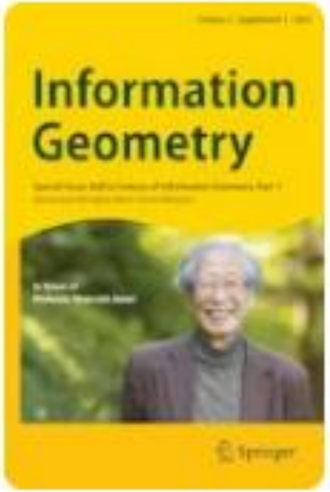
Independent component analysis in the light of Information Geometry

Jean-François Cardoso[1]

Received: 21 September 2022 / Revised: 10 October 2022 / Accepted: 12 October 2022 /
Published online: 24 October 2022

Abstract
I recall my first encounter with Professor Shun-ichi Amari who, once upon a time in Las Vegas, gave me a precious hint about connecting independent component analysis (ICA) to Information Geometry. The paper sketches, rather informally, some of the insights gained in following this lead.

Keywords Independent component analysis · Information Geometry · Pythagoras theorem · Non Gaussianity



Volume 7
Special Issue 1
Article 2
Number 94

SURVEY PAPER

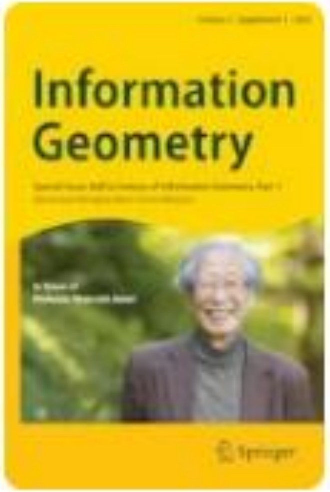
Geometry of EM and related iterative algorithms

Hideitsu Hino[1,2] · **Shotaro Akaho**[3] · **Noboru Murata**[4]

Received: 29 August 2022 / Revised: 12 November 2022 / Accepted: 13 November 2022 /
Published online: 19 November 2022

Abstract
The Expectation–Maximization (EM) algorithm is a simple meta-algorithm that has been used for many years as a methodology for statistical inference when there are missing measurements in the observed data or when the data is composed of observables and unobservables. Its general properties are well studied, and also, there are countless ways to apply it to individual problems. In this paper, we introduce the *em* algorithm, an information geometric formulation of the EM algorithm, and its extensions and applications to various problems. Specifically, we will see that it is possible to formulate an outlier–robust inference algorithm, an algorithm for calculating channel capacity, parameter estimation methods on probability simplex, particular multivariate analysis methods such as principal component analysis in a space of probability models and modal regression, matrix factorization, and learning generative models, which have recently attracted attention in deep learning, from the geometric perspective provided by Amari.

Keywords Information geometry · EM algorithm · *em* algorithm · Bregman divergence · Information theory · Robust statistics · Generative models



Volume 7
Special Issue 1
Article 3
Number 95

SURVEY PAPER

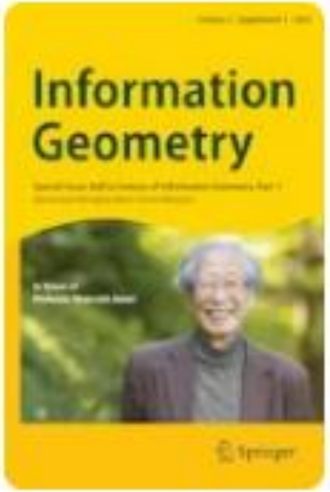
Hommage to Chentsov's theorem

Akio Fujiwara[1]

Received: 26 September 2022 / Revised: 6 November 2022 / Accepted: 12 November 2022 /
Published online: 21 November 2022

Abstract
Chentsov's theorem, which characterises Markov invariant Riemannian metric and affine connections of manifolds of probability distributions on finite sample spaces, is undoubtedly a cornerstone of information geometry. This article aims at providing a comprehensible survey of Chentsov's theorem as well as its modest extensions to generic tensor fields and to parametric models comprising continuous probability densities on $\mathbb{R}^k$.

Keywords Chentsov's theorem · Markov invariance · Fisher metric · α-connections · Amari–Chentsov tensor · Information geometry



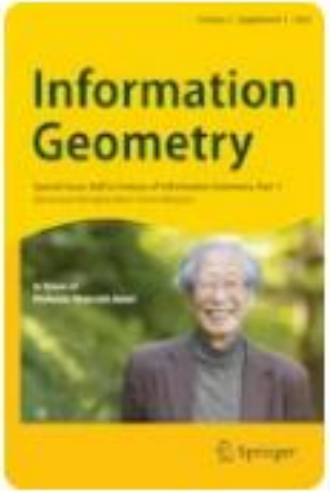
Volume 7
Special Issue 1
Article 4
Number 96

PERSPECTIVE

Toward differential geometry of statistical submanifolds

Hitoshi Furuhata[1]

Received: 31 August 2022 / Revised: 30 October 2022 / Accepted: 3 November 2022 /
Published online: 22 November 2022



Volume 7
Special Issue 1
Article 5
Number 97

Abstract

A brief introduction of doubly minimal submanifolds of statistical manifolds is given. A complex submanifold of a holomorphic statistical manifold is doubly minimal. Similar properties are obtained in the case where the ambient space is a Sasakian statistical manifold.

Keywords Statistical submanifolds · Doubly minimal · Doubly totally umbilical

SURVEY PAPER

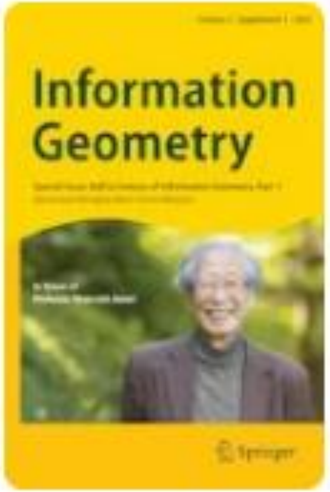
Affine statistical bundle modeled on a Gaussian Orlicz–Sobolev space

Giovanni Pistone[1]

Received: 14 October 2022 / Revised: 10 November 2022 / Accepted: 12 November 2022 /
Published online: 24 November 2022

Abstract
The dually flat structure of statistical manifolds can be derived in a non-parametric way from a particular case of affine space defined on a qualified set of probability measures. The statistically natural displacement mapping of the affine space depends on the notion of Fisher's score. The model space must be carefully defined if the state space is not finite. Among various options, we discuss how to use Orlicz–Sobolev spaces with Gaussian weight. Such a fully non-parametric set-up provides tools to discuss intrinsically infinite-dimensional evolution problems

Keywords Information geometry · Gaussian Orlicz–Sobolev space · Statistical bundle · Exponential manifold · Dually flat affine manifold



Volume 7
Special Issue 1
Article 6
Number 98

SURVEY PAPER

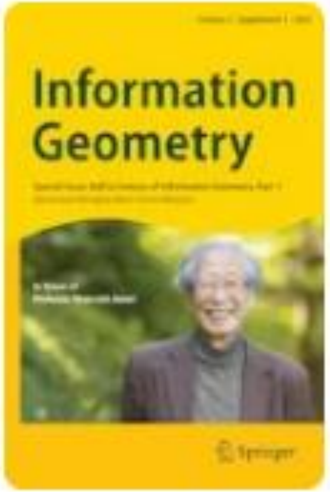
Conditional inference of Poisson models and information geometry: an ancillary review

Tomonari Sei[1]

Received: 31 August 2022 / Revised: 5 November 2022 / Accepted: 15 November 2022 / Published online: 27 November 2022

Abstract
The Poisson distribution is a fundamental tool in categorical data analysis. This paper reviews conditional inference for the independent Poisson model. It is noted that the conditioning variable is not an ancillary statistic in the exact sense except in the case of the product multinomial sampling scheme, whereas two versions of the ancillary property hold in general. The ancillary properties justify the use of conditional inference, as first proposed by R. A. Fisher and subsequently discussed by many researchers. The mixed coordinate system developed in information geometry is emphasized as effective for the description of facts.

Keywords A-hypergeometric distribution · Ancillary statistic · Asymptotic theory · Fisher information · Mixed coordinate system



Volume 7
Special Issue 1
Article 7
Number 99

RESEARCH PAPER

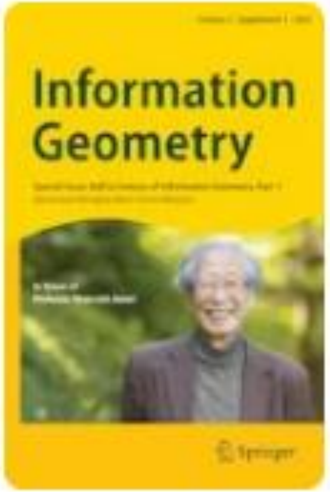
Generalized estimators, slope, efficiency, and fisher information bounds

Paul Vos[1]

Received: 10 August 2022 / Revised: 22 November 2022 / Accepted: 22 November 2022 /
Published online: 4 December 2022

Abstract
Point estimators may not exist, need not be unique, and their distributions are not parameter invariant. Generalized estimators provide distributions that are parameter invariant, unique, and exist when point estimates do not. Comparing point estimators using variance is less useful when estimators are biased. A squared slope Λ is defined that can be used to compare both point and generalized estimators and is unaffected by bias. Fisher information I and variance are fundamentally different quantities: the latter is defined at a distribution that need not belong to a family, while the former cannot be defined without a family of distributions, M. Fisher information and Λ are similar quantities as both are defined on the tangent bundle TM and I provides an upper bound, $\Lambda \leq I$, that holds for all sample sizes—asymptotics are not required. Comparing estimators using Λ rather than variance supports Fisher's claim that I provides a bound even in small samples. Λ-efficiency is defined that extends the efficiency of unbiased estimators based on variance. While defined by the slope, Λ-efficiency is simply ρ^2, the square of the correlation between estimator and score function.

Keywords Nonasymptotics · Generalized estimators · Slope · Fisher information · Efficiency



Volume 7
Special Issue 1
Article 8
Number 100

RESEARCH PAPER

A two-parameter family of non-parametric, deformed exponential manifolds

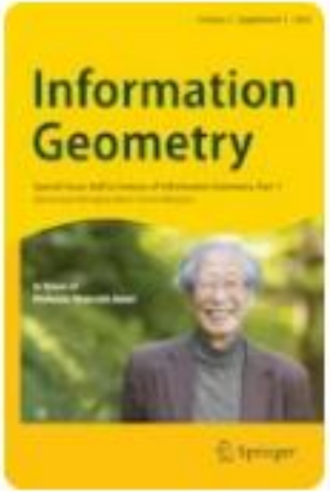
Nigel J. Newton[1]

Received: 1 September 2022 / Revised: 1 September 2022 / Accepted: 13 November 2022 /
Published online: 6 December 2022

Abstract

We construct a new family of non-parametric statistical manifolds by means of a two-parameter class of deformed exponential functions, that includes functions with power-law, linear and sublinear rates of growth. The manifolds are modelled on weighted, mixed-norm Sobolev spaces that are especially suited to this purpose, in the sense that an important class of nonlinear superposition operators (those used in the construction of divergences and tensors) act continuously on them. We analyse variants of these operators, that map into "subordinate" Sobolev spaces, and evaluate the associated gain in regularity. With appropriate choice of parameter values, the manifolds support a large variety of the statistical divergences and entropies appearing in the literature, as well as their associated tensors, eg. the Fisher-Rao metric. Manifolds of finite measures and probability measures are constructed; the latter are shown to be smoothly embedded submanifolds of the former.

Keywords Banach manifold · Fisher-Rao metric · Information Theory · Log-Sobolev inequality · Non-parametric statistics · Sobolev spaces



Volume 7
Special Issue 1
Article 9
Number 101

SURVEY PAPER

Recent advances in algebraic geometry and Bayesian statistics

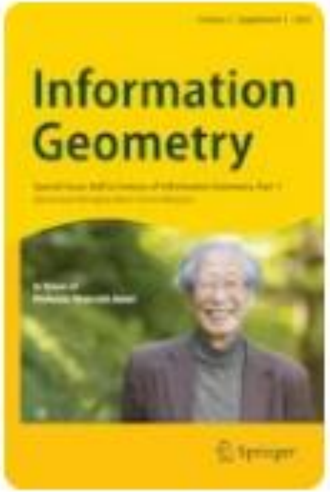
Sumio Watanabe[1]

Received: 30 August 2022 / Revised: 2 November 2022 / Accepted: 18 November 2022 /
Published online: 6 December 2022

Abstract

This article is a review of theoretical advances in the research field of algebraic geometry and Bayesian statistics in the last two decades. Many statistical models and learning machines which contain hierarchical structures or latent variables are called nonidentifiable, because the map from a parameter to a statistical model is not one-to-one. In nonidentifiable models, both the likelihood function and the posterior distribution have singularities in general, hence it was difficult to analyze their statistical properties. However, from the end of the 20th century, new theory and methodology based on algebraic geometry have been established which enable us to investigate such models and machines in the real world. In this article, the following results in recent advances are reported. First, we explain the framework of Bayesian statistics and introduce a new perspective from the birational geometry. Second, two mathematical solutions are derived based on algebraic geometry. An appropriate parameter space can be found by a resolution map, which makes the posterior distribution be normal crossing and the log likelihood ratio function be well-defined. Third, three applications to statistics are introduced. The posterior distribution is represented by the renormalized form, the asymptotic free energy is derived, and the universal formula among the generalization loss, the cross validation, and the information criterion is established. Two mathematical solutions and three applications to statistics based on algebraic geometry reported in this article are now being used in many practical fields in data science and artificial intelligence.

Keywords Birational geometry · Resolution of singularities · Bayesian statistics · Real log canonical threshold



Volume 7
Special Issue 1
Article 10
Number 102

SURVEY PAPER

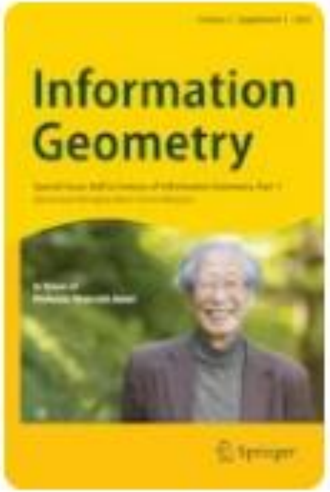
Geometry and applied statistics

Paul Marriott[1]

Received: 5 September 2022 / Revised: 25 November 2022 / Accepted: 26 November 2022 /
Published online: 6 December 2022

Abstract
We take a very high level overview of the relationship between Geometry and Applied Statistics 50 years from the birth of Information Geometry. From that date we look both backwards and forwards. We show that Geometry has always been part of the statistician's toolbox and how it played a vital role in the evolution of Statistics in the last 50 years.

Keywords Applied statistics · Information geometry



Volume 7
Special Issue 1
Article 11
Number 103

RESEARCH PAPER

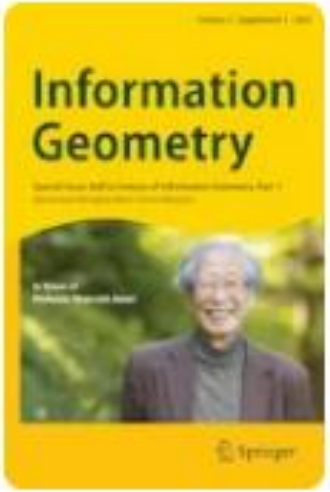
Minimum information divergence of Q-functions for dynamic treatment resumes

Shinto Eguchi[1]

Received: 3 September 2022 / Revised: 17 November 2022 / Accepted: 18 November 2022 /
Published online: 9 December 2022

Abstract
This paper aims at presenting a new application of information geometry to reinforcement learning focusing on dynamic treatment resumes. In a standard framework of reinforcement learning, a Q-function is defined as the conditional expectation of a reward given a state and an action for a single-stage situation. We introduce an equivalence relation, called the policy equivalence, in the space of all the Q-functions. A class of information divergence is defined in the Q-function space for every stage. The main objective is to propose an estimator of the optimal policy function by a method of minimum information divergence based on a dataset of trajectories. In particular, we discuss the γ-power divergence that is shown to have an advantageous property such that the γ-power divergence between policy-equivalent Q-functions vanishes. This property essentially works to seek the optimal policy, which is discussed in a framework of a semiparametric model for the Q-function. The specific choices of power index γ give interesting relationships of the value function, and the geometric and harmonic means of the Q-function. A numerical experiment demonstrates the performance of the minimum γ-power divergence method in the context of dynamic treatment regimes.

Keywords Information geometry · Geometric mean · Minimum divergence · Optimal policy · Semiparametric model · Value function



Volume 7
Special Issue 1
Article 12
Number 104

RESEARCH PAPER

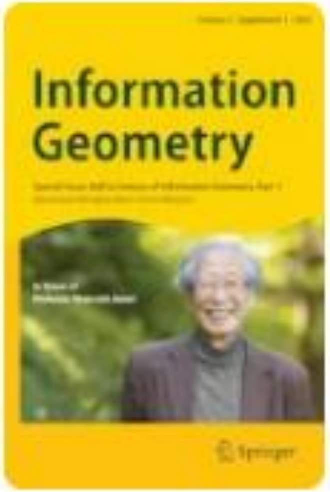
Hierarchy of deformations in concavity

Kazuhiro Ishige[1] · Paolo Salani[2] · Asuka Takatsu[3,4]

Received: 1 September 2022 / Revised: 25 November 2022 / Accepted: 28 November 2022 /
Published online: 12 December 2022

Abstract
A *deformation* is a positive continuous function defined on an appropriate interval. Through deformations, we generalize the notion of concavity for functions. We introduce the *order function* of a deformation, which permits to determine precisely the ranking of a deformation by taking account of the corresponding concavities. In the hierarchy, the action of positive constant multiples provides an equivalence relation and, if we focus on C^1-deformations, a one-to-one correspondence between the equivalence classes and the order functions is determined. Deformations having a constant valued order function play a fundamental role, and this is only the case of power functions. We show that the concavity associated to a deformation whose order function is nonincreasing and uniformly bounded from above by $q \in \mathbb{R}$ can approximate the concavity associated to the power function of exponent q. Finally, we review three examples of deformations whose order function is nonincreasing and uniformly bounded from above. One is a power function, and the others are related to a concavity preserved by the Dirichlet heat flow in convex domains of Euclidean space.

Keywords Deformation · Deformed logarithmic function · Generalized concavity · Hierarchy



Volume 7
Special Issue 1
Article 13
Number 105

SURVEY PAPER

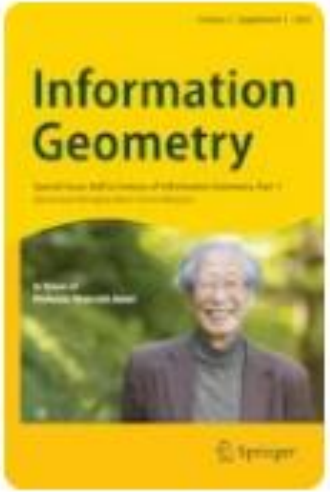
Natural differentiable structures on statistical models and the Fisher metric

Hông Vân Lê[1]

Received: 24 August 2022 / Revised: 20 November 2022 / Accepted: 5 December 2022 /
Published online: 13 December 2022

Abstract
In this paper I discuss the relation between the concept of the Fisher metric and the concept of differentiability of a family of probability measures. I compare the concepts of smooth statistical manifolds, differentiable families of measures, k-integrable parameterized measure models, diffeological statistical models, differentiable measures, which arise in Information Geometry, mathematical statistics and measure theory, and discuss some related problems.

Keywords Smooth statistical model · Differentiable family of measures · Fisher–Rao metric · Amari–Chentsov tensor · Parameterized measure model · Diffeological statistical model · Cramér–Rao inequality



Volume 7
Special Issue 1
Article 14
Number 106

PERSPECTIVE

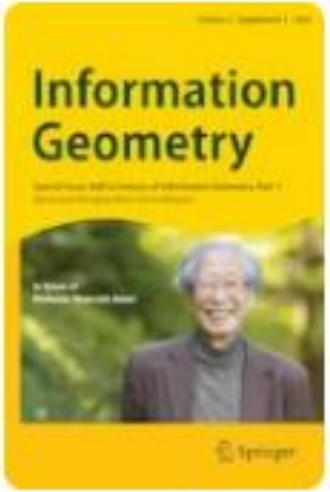
Uncertainty and Quantum Variance at the light of Quantum Information Geometry

Paolo Gibilisco[1]

Received: 30 October 2022 / Revised: 22 November 2022 / Accepted: 28 November 2022 /
Published online: 14 December 2022

Abstract
Using the Kubo-Ando operator means Denes Petz was able to explain what Quantum Covariance(s) and Quantum Fisher Information(s) are. In this paper the relation between QC and QFI is explained and some of its consequences, notably the Dynamical Uncertainty Principle, are described. Finally some new directions for Classical and Quantum Information Geometry are suggested.

Keywords Quantum Covariance · Quantum Fisher Information · Metric Adjusted Skew Information · Dynamical Uncertainty Principle



Volume 7
Special Issue 1
Article 15
Number 107

RESEARCH PAPER

Conformal mirror descent with logarithmic divergences

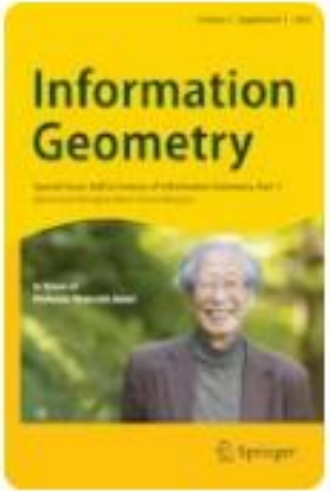
Amanjit Singh Kainth[1,3] · Ting-Kam Leonard Wong[2] · Frank Rudzicz[1,3]

Received: 1 September 2022 / Revised: 21 November 2022 / Accepted: 30 November 2022 /
Published online: 14 December 2022

Abstract

The logarithmic divergence is an extension of the Bregman divergence motivated by optimal transport and a generalized convex duality, and satisfies many remarkable properties. Using the geometry induced by the logarithmic divergence, we introduce a generalization of continuous time mirror descent that we term the conformal mirror descent. We derive its dynamics under a generalized mirror map, and show that it is a time change of a corresponding Hessian gradient flow. We also prove convergence results in continuous time. We apply the conformal mirror descent to online estimation of a generalized exponential family, and construct a family of gradient flows on the unit simplex via the Dirichlet optimal transport problem.

Keywords Mirror descent · Gradient flow · Logarithmic divergence · Conformal Hessian metric · λ-duality · λ-exponential family



Volume 7
Special Issue 1
Article 16
Number 107

RESEARCH PAPER

Parametric models and information geometry on W*-algebras

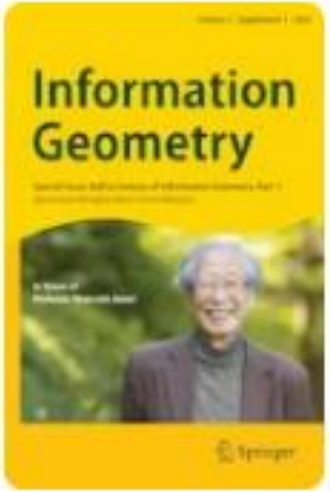
F. M. Ciaglia[1] · F. Di Nocera[2] · J. Jost[2] · L. Schwachhöfer[3]

Received: 5 September 2022 / Revised: 20 December 2022 / Accepted: 21 December 2022 /
Published online: 12 January 2023

Abstract

We introduce the notion of smooth parametric model of normal positive linear functionals on possibly infinite-dimensional $W^\star$-algebras generalizing the notions of parametric models used in classical and quantum information geometry. We then use the Jordan product naturally available in this context in order to define a Riemannian metric tensor on parametric models satsfying suitable regularity conditions. This Riemannian metric tensor reduces to the Fisher–Rao metric tensor, or to the Fubini-Study metric tensor, or to the Bures–Helstrom metric tensor when suitable choices for the $W^\star$-algebra and the models are made.

Keywords Probability distributions · Quantum states · C^*-algebras · Banach manifolds



Volume 7
Special Issue 1
Article 17
Number 108

RESEARCH PAPER

Gram matrices of quantum channels via quantum Fisher information with applications to decoherence and uncertainty

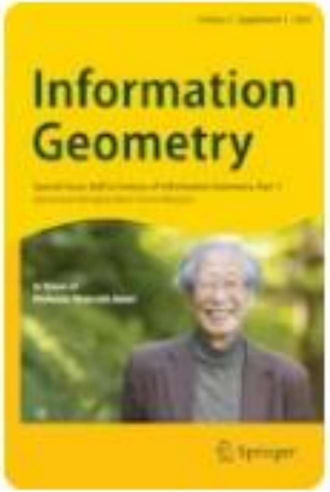
Shunlong Luo[1,2] · **Yuan Sun**[3]

Received: 22 August 2022 / Revised: 21 December 2022 / Accepted: 5 January 2023 /
Published online: 22 January 2023

Abstract

By use of Gram matrices associated with various versions of quantum Fisher information (related to monotone metrics and metric-adjusted skew information), we study information and geometry of quantum channels by introducing several quantities involving commutators and anticommutators for operators and quantum states. We reveal their basic features and evaluate these quantities for some important channels. As applications, we employ Gram matrices to quantify decoherence of quantum channels caused on quantum states, and establish some uncertainty relations, which refine the conventional Heisenberg uncertainty relations involving variance.

Keywords Quantum channels · Quantum Fisher information · Commutator · Anticommutator · Gram matrices



Volume 7
Special Issue 1
Article 18
Number 109

RESEARCH PAPER

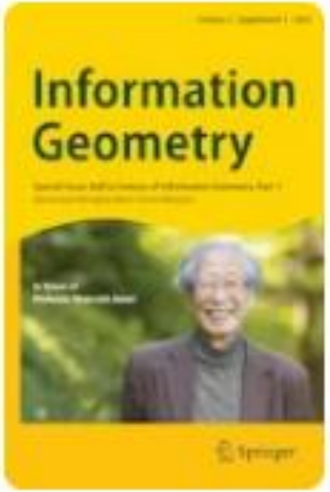
The exponential Orlicz space in quantum information geometry

Anna Jenčová[1]

Received: 10 October 2022 / Revised: 2 January 2023 / Accepted: 5 January 2023 /
Published online: 24 January 2023

Abstract
We review the construction of a quantum version of the exponential statistical manifold over the set of all faithful normal positive functionals on a von Neumann algebra. The construction is based on the relative entropy approach to state perturbation. We construct a quantum version of the exponential Orlicz space and discuss the properties of this space and its dual with respect to Kosaki L_p-spaces. We show that the constructed manifold admits a canonical divergence satisfying a Pythagorean relation. We also prove that the manifold structure is invariant under sufficient channels.

Keywords Quantum exponential manifold · Quantum relative entropy · Perturbation of states · Canonical divergence



Volume 7
Special Issue 1
Article 19
Number 110

RESEARCH PAPER

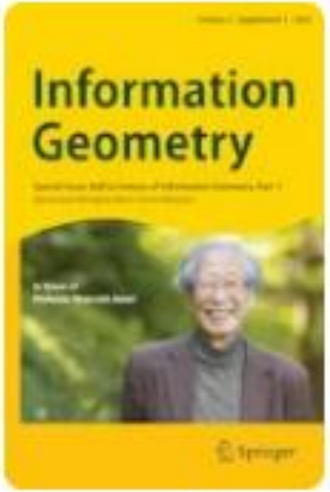
Projections of SDEs onto submanifolds

John Armstrong[1] · Damiano Brigo[2] · Emilio Ferrucci[3]

Received: 2 September 2022 / Revised: 13 December 2022 / Accepted: 20 December 2022 /
Published online: 2 February 2023

Abstract
In Armstrong et al. (Proc Lond Math Soc (3) 119(1):176–213, 2019) the authors define three projections of $\mathbb{R}^d$-valued stochastic differential equations (SDEs) onto submanifolds: the Stratonovich, Itô-vector and Itô-jet projections. In this paper, after a brief survey of SDEs on manifolds, we begin by giving these projections a natural, coordinate-free description, each in terms of a specific representation of manifold-valued SDEs. We proceed by deriving formulae for the three projections in ambient $\mathbb{R}^d$-coordinates. We use these to show that the Itô-vector and Itô-jet projections satisfy respectively a weak and mean-square optimality criterion "for small t": this is achieved by solving constrained optimisation problems. These results confirm, but do not rely on the approach taken in Armstrong et al. (Proc Lond Math Soc (3) 119(1):176–213, 2019), which is formulated in terms of weak and strong Itô–Taylor expansions. In the final section we exhibit examples showing how the three projections can differ, and explore alternative notions of optimality.

Keywords Ito · Stratonovich · Projection



Volume 7
Special Issue 1
Article 20
Number 111

RESEARCH PAPER

A parameterisation-invariant modification of the score test

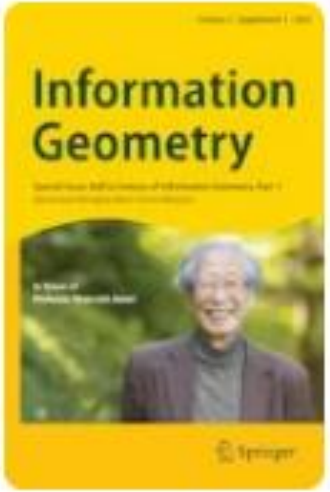
P. E. Jupp[1]

Received: 31 August 2022 / Revised: 10 December 2022 / Accepted: 4 February 2023 /
Published online: 7 March 2023

Abstract

The null distribution of the score test statistic is asymptotically chi-squared for large samples. The error in this approximation is improved greatly by a cubic modification. The coefficients of this cubic that are given in the literature depend on the parameterisation. This paper provides parameterisation-invariant versions of the coefficients, expresses them in terms of appropriate tensors, and provides geometric interpretations.

Keywords Bartlett correction · Generalised Bartlett correction · Interest parameter · Invariant Taylor expansion · Large-sample asymptotics · Likelihood yoke · Tensor



Volume 7
Special Issue 1
Article 21
Number 112

RESEARCH PAPER

Geometric thermodynamics for the Fokker–Planck equation: stochastic thermodynamic links between information geometry and optimal transport

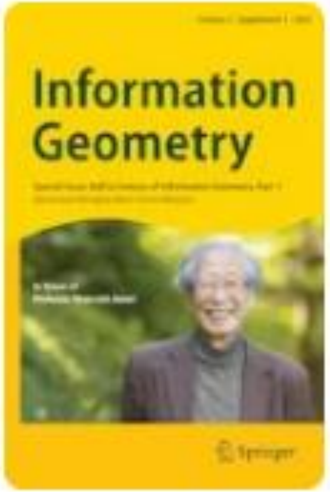
Sosuke Ito[1]

Received: 7 September 2022 / Revised: 31 January 2023 / Accepted: 6 February 2023 /
Published online: 9 March 2023

Abstract

We propose a geometric theory of non-equilibrium thermodynamics, namely geometric thermodynamics, using our recent developments of differential-geometric aspects of entropy production rate in non-equilibrium thermodynamics. By revisiting our recent results on geometrical aspects of entropy production rate in stochastic thermodynamics for the Fokker–Planck equation, we introduce a geometric framework of non-equilibrium thermodynamics in terms of information geometry and optimal transport theory. We show that the proposed geometric framework is useful for obtaining several non-equilibrium thermodynamic relations, such as thermodynamic trade-off relations between the thermodynamic cost and the fluctuation of the observable, optimal protocols for the minimum thermodynamic cost and the decomposition of the entropy production rate for the non-equilibrium system. We clarify several stochastic-thermodynamic links between information geometry and optimal transport theory via the excess entropy production rate based on a relation between the gradient flow expression and information geometry in the space of probability densities and a relation between the velocity field in optimal transport and information geometry in the space of path probability densities.

Keywords Stochastic thermodynamics · Entropy production · Information geometry · Optimal transport theory · Fokker–Planck equation



Volume 7
Special Issue 1
Article 22
Number 113

RESEARCH PAPER

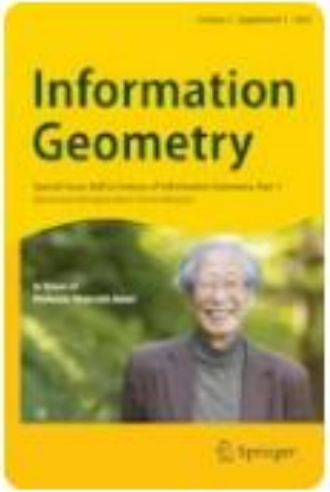
Geometric thermodynamics for the Fokker–Planck equation: stochastic thermodynamic links between information geometry and optimal transport

Sosuke Ito[1]

Received: 7 September 2022 / Revised: 31 January 2023 / Accepted: 6 February 2023 /
Published online: 9 March 2023

Abstract
We propose a geometric theory of non-equilibrium thermodynamics, namely geometric thermodynamics, using our recent developments of differential-geometric aspects of entropy production rate in non-equilibrium thermodynamics. By revisiting our recent results on geometrical aspects of entropy production rate in stochastic thermodynamics for the Fokker–Planck equation, we introduce a geometric framework of non-equilibrium thermodynamics in terms of information geometry and optimal transport theory. We show that the proposed geometric framework is useful for obtaining several non-equilibrium thermodynamic relations, such as thermodynamic trade-off relations between the thermodynamic cost and the fluctuation of the observable, optimal protocols for the minimum thermodynamic cost and the decomposition of the entropy production rate for the non-equilibrium system. We clarify several stochastic-thermodynamic links between information geometry and optimal transport theory via the excess entropy production rate based on a relation between the gradient flow expression and information geometry in the space of probability densities and a relation between the velocity field in optimal transport and information geometry in the space of path probability densities.

Keywords Stochastic thermodynamics · Entropy production · Information geometry · Optimal transport theory · Fokker–Planck equation



Volume 7
Special Issue 1
Article 23
Number 114

SURVEY PAPER

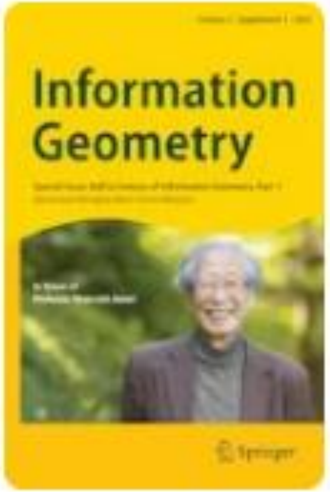
Optimal projection filters with information geometry

John Armstrong[1] · Damiano Brigo[2] · Bernard Hanzon[3]

Received: 18 February 2023 / Revised: 3 May 2023 / Accepted: 13 May 2023 /
Published online: 2 June 2023

Abstract
We review the introduction of several types of projection filters. Projection structures coming from information geometry are used to obtain a finite dimensional filter in the form of a stochastic differential equation (SDE), starting from the exact infinite-dimensional stochastic partial differential equation (SPDE) for the optimal filter. We start with the Stratonovich projection filters based on the Hellinger distance as introduced and developed in Brigo et al. (IEEE Trans Autom Control 43(2):247–252, 1998, Bernoulli 5(3):495–534, 1999), where the SPDE is put in Stratonovich form before projection, hence the term "Stratonovich projection". The correction step of the filtering algorithm can be made exact by choosing a suitable exponential family as manifold, there is equivalence with assumed density filters and numerical examples have been studied. Other authors further developed these projection filters and we present a brief literature review. A second type of Stratonovich projection filters was introduced in Armstrong and Brigo (Math Control Signals Syst 28(1):1–33, 2016) where a direct L^2 metric is used for projection. Projecting on mixtures of densities as a manifold coincides with Galerkin methods. All the above projection filters lack optimality, as the single vector fields of the Stratonovich SPDE are projected optimally but the SPDE solution as a whole is not approximated optimally by the projected SDE solution according to a clear criterion. This led to the optimal projection filters in Armstrong et al. (Proc Lond Math Soc 119(1):176–213, 2019, Projection of SDEs onto submanifolds. "Information Geometry", 2023 special issue on half a century of information geometry, 2018), based on the Ito vector and Ito jet projections, where several types of mean square distances between the optimal filter SPDE solution and the sought finite dimensional SDE approximations are minimized, with numerical examples. After reviewing the above developments, we conclude with the remaining challenges.

Keywords Stochastic partial differential equations · Stochastic differential equations · SPDEs projection on a submanifold · Stratonovich projection · Itô-vector projection · Itô-jet projection · Nonlinear filtering · Projection filters · Stratonovich projection filters · Optimal projection filters



Volume 7
Special Issue 1
Article 24
Number 115

RESEARCH PAPER

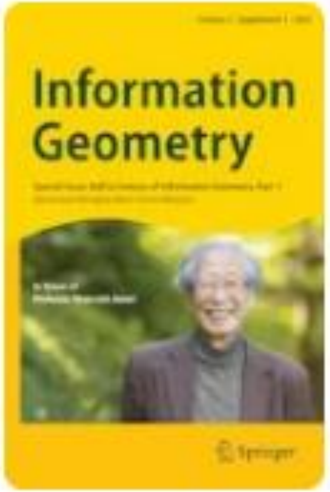
A certain ODE-system defining the geometric divergence

Takashi Kurose[1]

Received: 13 December 2022 / Revised: 27 April 2023 / Accepted: 25 May 2023 /
Published online: 5 June 2023

Abstract
A system of ordinary differential equations satisfied by the geometric divergence, a generalization of the canonical divergence on a dually flat space, is introduced. By using the system, it is shown that a family of contrast functions can be constructed on an arbitrary statistical manifold in a certain canonical manner.

Keywords Statistical manifolds · Contrast functions · Geometric divergences



Volume 7
Special Issue 1
Article 25
Number 116

RESEARCH PAPER

Doubly autoparallel structure and curvature integrals

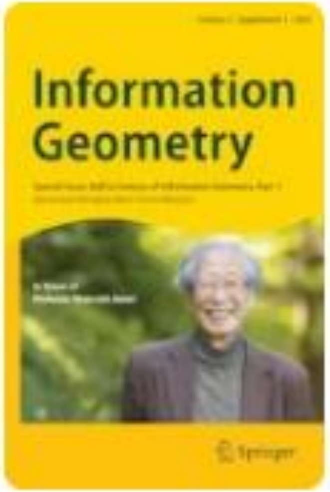
Applications to iteration complexity for solving convex programs

Atsumi Ohara[1] · Hideyuki Ishi[2] · Takashi Tsuchiya[3]

Received: 9 September 2022 / Revised: 4 July 2023 / Accepted: 9 July 2023 /
Published online: 27 July 2023

Abstract
On a statistical manifold, we can define autoparallel submanifolds and path integrals of the second fundamental forms (*curvature integrals*) for its primal and dual affine connections, respectively. A submanifold is called *doubly autoparallel* if it is simultaneously autoparallel with respect to the both connections. In this paper we first discuss common properties of such submanifolds. In particular we next give an algebraic characterization of them in Jordan algebras and show their applications. Further, we exhibit that both curvature integrals induced from dually flat structure are interestingly related to an unexpected quantity, i.e., iteration-complexity of the interior-point algorithms for convex optimization defined on a submanifold that is *not* doubly autoparallel.

Keywords Doubly autoparallel structure · Jordan algebra · Curvature integral · Interior point method · Structured covariance estimation



Volume 7
Special Issue 1
Article 26
Number 117

RESEARCH PAPER

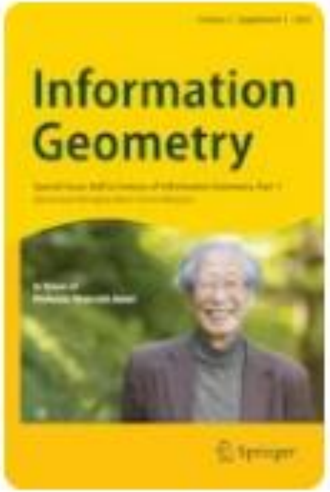
G-dual Teleparallel Connections in Information Geometry

F. M. Ciaglia[1] · F. Di Cosmo[1,2] · A. Ibort[1,2] · G. Marmo[3,4]

Received: 26 July 2022 / Revised: 13 June 2023 / Accepted: 26 July 2023 /
Published online: 25 August 2023

Abstract
Given a real, finite-dimensional, smooth parallelizable Riemannian manifold $(\mathcal{N}, G)$ endowed with a teleparallel connection ∇ determined by a choice of a global basis of vector fields on $\mathcal{N}$, we show that the G-dual connection ∇^* of ∇ in the sense of Information Geometry must be the teleparallel connection determined by the basis of G-gradient vector fields associated with a basis of differential one-forms which is (almost) dual to the basis of vector fields determining ∇. We call any such pair (∇, ∇^*) a ***G*-dual teleparallel pair**. Then, after defining a covariant $(0, 3)$ tensor T uniquely determined by $(\mathcal{N}, G, \nabla, \nabla^*)$, we show that T being symmetric in the first two entries is equivalent to ∇ being torsion-free, that T being symmetric in the first and third entry is equivalent to ∇^* being torsion free, and that T being symmetric in the second and third entries is equivalent to the basis vectors determining ∇ (∇^*) being parallel-transported by ∇^* (∇). Therefore, G-dual teleparallel pairs provide a generalization of the notion of Statistical Manifolds usually employed in Information Geometry, and we present explicit examples of G-dual teleparallel pairs arising both in the context of both Classical and Quantum Information Geometry.

Keywords Probability distributions · Quantum states · Dual connections



Volume 7
Special Issue 1
Article 27
Number 118

RESEARCH PAPER

Estimation with infinite-dimensional exponential family and Fisher divergence

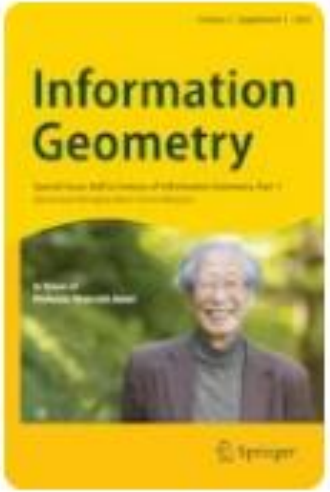
Kenji Fukumizu[1]

Received: 28 February 2023 / Revised: 25 October 2023 / Accepted: 29 October 2023 /
Published online: 13 November 2023

Abstract

Infinite dimensional exponential families have been theoretically studied, but their practical applications are still limited because empirical estimation is not straightforward. This paper first gives a brief survey of studies on the estimation method for infinite-dimensional exponential families. The method uses score matching, which is based on the Fisher divergence. The second topic is to investigate the Fisher divergence as a member of an extended family of divergences, which employ operators in defining divergences.

Keywords Infinite-dimensional exponential family · Fisher divergence · Score matching



Volume 7
Special Issue 1
Article 28
Number 119

SURVEY PAPER

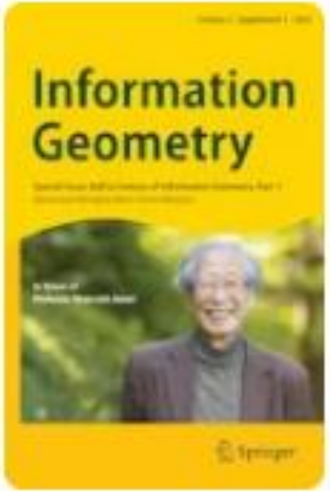
Legendre duality: from thermodynamics to information geometry

Jan Naudts[1] · Jun Zhang[2]

Received: 29 March 2023 / Revised: 19 October 2023 / Accepted: 27 October 2023 /
Published online: 8 November 2023

Abstract
This paper reviews the role of convex duality in Information Geometry. It clarifies the notion of bi-orthogonal coordinates associated with Legendre duality by treating its two underlying aspects separately: as a dual coordinate system and as a bi-orthogonal frame. It addresses the deformation of exponential families in a way that stills preserves the dually-flat geometry of 1- and (-1)-connections. The deformation involves a metric which generalizes the Fisher–Rao metric controlled by one degree of freedom and a pair of connections controlled by an additional degree of freedom.

Keywords Convex duality · Dually-flat geometry · Deformed exponential and logarithmic functions · Rho–tau connections



Volume 7
Special Issue 1
Article 29
Number 120

RESEARCH PAPER

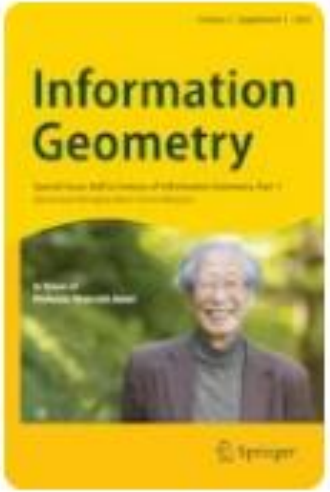
The Fisher metric as a metric on the cotangent bundle

Hiroshi Nagaoka[1]

Received: 11 September 2023 / Revised: 10 November 2023 / Accepted: 12 November 2023 / Published online: 21 December 2023

Abstract
The Fisher metric on a manifold of probability distributions is usually treated as a metric on the tangent bundle. In this paper, we focus on the metric on the cotangent bundle induced from the Fisher metric with calling it the Fisher co-metric. We show that the Fisher co-metric can be defined directly without going through the Fisher metric by establishing a natural correspondence between cotangent vectors and random variables. This definition clarifies a close relation between the Fisher co-metric and the variance/covariance of random variables, whereby the Cramér-Rao inequality is trivialized. We also discuss the monotonicity and the invariance of the Fisher co-metric with respect to Markov maps, and present a theorem characterizing the co-metric by the invariance, which can be regarded as a cotangent version of Čencov's characterization theorem for the Fisher metric. The obtained theorem can also be viewed as giving a characterization of the variance/covariance.

Keywords Information geometry · Fisher metric · Cotangent space · Čencov's (Chentsov's) theorem



Volume 7
Special Issue 1
Article 30
Number 121