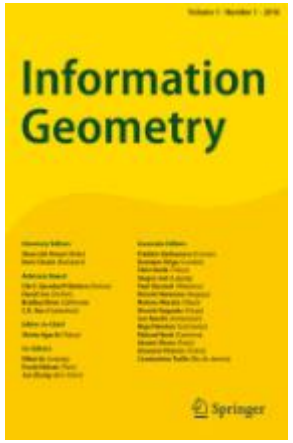
RESEARCH

Some computational results on Koszul–Vinberg cochain complexes

Hanwen Liu [1,3] · Jun Zhang[2,3]

Received: 28 April 2024 / Revised: 6 September 2024 / Accepted: 13 October 2024 /
Published online: 23 December 2024



Volume 8
Issue 1
Article 1
Number 131

Abstract

An affine connection is said to be flat if its curvature tensor vanishes identically. Koszul–Vinberg (KV for abbreviation) cohomology has been invoked to study the deformation theory of flat and torsion-free affine connections on tangent bundle. In this Note, we compute explicitly the differentials of various specific KV cochains, and study their relation to classical objects in information geometry, including deformations associated with projective and dual-projective transformations of a flat and torsion-free affine connection. As an application, we also give a simple yet non-trivial example of a left-symmetric algebra of which the second cohomology group does not vanish.

Keywords Affine structure · Koszul–Vinberg cochain complexes · Conformal and projective transform · Exterior covariant derivative

RESEARCH PAPER

Information geometry of system spaces

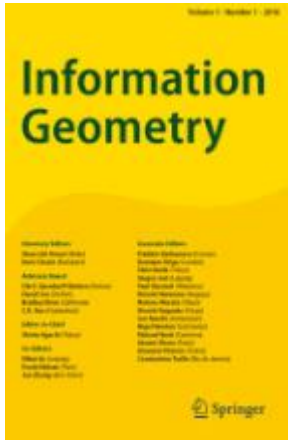
Masayuki Kumon[1]

Received: 16 September 2022 / Revised: 23 December 2024 / Accepted: 26 December 2024 /
Published online: 1 February 2025

Abstract

For the totality of square and regular transfer function matrices, basic differential geometrical structures such as Riemannian metrics and linear connections are introduced from the viewpoint of left and right invariance. These structures are successively introduced into the outer and inner system spaces, and then into the second and higher order cumulant spectral density spaces. The dualities of the derived geometrical structures with torsions and curvatures are investigated, further the possibilities for the existences of divergences are explored for the respective spaces. As an application of the information geometrical structures, we analyze the block oriented nonlinear output feedback mechanisms, and clarify the roles of the associated sensitivity operators. We also obtain the conditions with desired output feedback rules for which the divergences attain the ultimate zero values. Illustrative examples are given to the readers for the block oriented nonlinear output feedback expressions.

Keywords Block oriented model · Covariant derivative · Cumulant spectral density · Divergence · Information geometry · Lie group · Mutual information quantity · Nonlinear output feedback · Riemannian metric · Riemann–Christoffel curvature · Sensitivity operator · Torsion · Volterra series · Wiener series



Volume 8
Issue 1
Article 2
Number 132

RESEARCH

F-t joint distributions on real Siegel domains

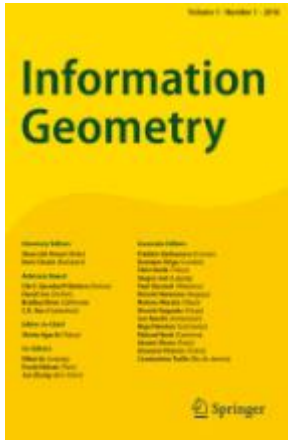
Hiroto Inoue[1]

Received: 31 July 2024 / Revised: 26 January 2025 / Accepted: 31 January 2025 /
Published online: 21 February 2025

Abstract

In statistics, the homogeneous cones are considered as the general parameter spaces for variance matrices related with the graphical models. For the argument of probability distributions, the Wishart distributions are defined based on the analysis of homogeneous cones, as the image of the normal distributions by quadratic maps. Now, the parameter space of pairs of variance matrix and mean vector is identified with more general homogeneous domain, called the real Siegel domain. In the present work, we consider random variables valued in the real Siegel domain associated with the quadratic map of homogeneous cone, and derive their probability density functions via the transformation group. This derivation is also based on the analysis on homogeneous cones including the gamma and beta functions due to Gindikin. We call our result the F-t joint distribution on the real Siegel domain, since its marginal distributions give an extension of the F and t-distributions in 1-dimensional case. Especially, it enables us to describe a joint estimation and a simultaneous hypothesis testing for the pair of parameters of normal distribution.

Keywords Homogeneous cone · Quadratic map · Wishart distribution · Simultaneous hypothesis testing



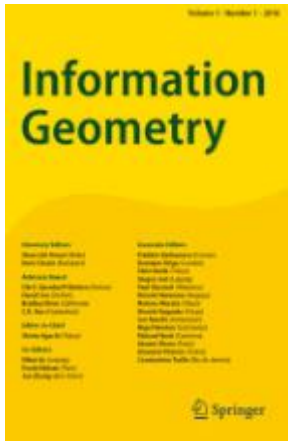
Volume 8
Issue 1
Article 3
Number 133

RESEARCH

Generalized estimation and information

Paul W. Vos[1] · Qiang Wu[1]

Received: 3 September 2024 / Revised: 17 December 2024 / Accepted: 9 February 2025 /
Published online: 22 February 2025



Volume 8
Issue 1
Article 4
Number 134

Abstract

This paper extends the idea of a generalized estimator for a scalar parameter (Vos in Inform Geom 7:151–170, 2022) to multi-dimensional parameters both with and without nuisance parameters. The title reflects the fact that generalized estimators provide more than simply another method to find point estimators, and that the methods to assess generalized estimators differ from those for point estimators. By *generalized estimation* we mean the use of generalized estimators together with an extended definition of *information* to assess their inferential properties. We show that Fisher information provides an upper bound for the information utilized by an estimator and that the score attains this bound. This optimality result provides theoretical justification for likelihood-based inference, effectively narrowing the search for optimal estimators when the score is computationally feasible.

Keywords Cramér–Rao bound · Fisher information · Geometry · Score · Slope

RESEARCH PAPER

Dual stochastic natural gradient descent

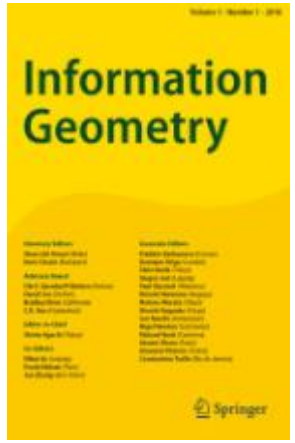
and convergence of interior half-space gradient approximations

Borja Sánchez-López[1] · Jesus Cerquides[1]

Received: 1 April 2020 / Revised: 13 December 2024 / Accepted: 2 March 2025 /
Published online: 4 April 2025

Abstract
The multinomial logistic regression (MLR) model is widely used in statistics and machine learning. On the one hand, stochastic gradient descent (SGD) is the most common approach for determining the parameters of a such model in big data scenarios, due to its simplicity and low computational complexity property. Furthermore, SGD has proven convergence under reasonable conditions. However, SGD has slow sublinear rates of convergence and it often reduces convergence speed due to the plateau phenomenon. On the other hand, stochastic natural gradient descent (SNGD), proposed by Amari, is a manifold optimization method shown to be Fisher efficient when it converges, but its convergence properties remain unproven and it is often computationally prohibitive for models with a large number of parameters. Here, we propose dual stochastic natural gradient descent (DSNGD), a stochastic optimization method for MLR based on manifold optimization concepts. In the discrete scenario, DSNGD (i) has linear per-iteration computational complexity in the number of parameters, and (ii) is proven to converge. To achieve (i) we leverage the dual flatness of the family of joint distributions for MLR to simplify computations. To ensure (ii) DSNGD builds on the foundational ideas of convergent stochastic natural gradient descent (CSNGD), a variant of SNGD with guaranteed convergence, using an independent sequence to construct a bounded approximation of the natural gradient. By generalizing a result from Sunehag et al., we prove that DSNGD converges in the discrete case and maintains linear computational complexity per iteration. Beyond its convergence property and linear computational complexity, DSNGD empirically demonstrates fast convergence comparable to SNGD, improves upon SGD performance, and exhibits stability where SNGD does not.

Keywords Multinomial logistic regression · Stochastic gradient descent · Natural gradient · Convergence · Riemannian manifold · Computational complexity



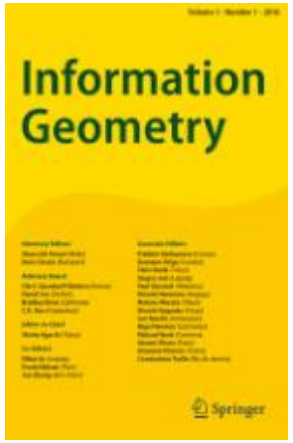
Volume 8
Issue 1
Article 5
Number 135

RESEARCH

Maximum information divergence from linear and toric models

Yulia Alexandr[1] · Serkan Hoşten[2]

Received: 24 April 2024 / Revised: 15 March 2025 / Accepted: 30 March 2025 /
Published online: 21 April 2025



Volume 8
Issue 1
Article 6
Number 136

Abstract

We study the problem of maximizing information divergence from a new perspective using logarithmic Voronoi polytopes. We show that for linear models, the maximum is always achieved at the boundary of the probability simplex. For toric models, we present an algorithm that combines the combinatorics of the chamber complex with numerical algebraic geometry. We pay special attention to reducible models and models of maximum likelihood degree one.

Keywords KL divergence · Information divergence · Maximum likelihood estimation · Toric models