RESEARCH PAPER

Wasserstein Riemannian geometry of Gaussian densities

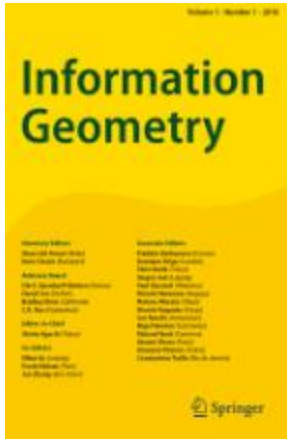
Luigi Malagò[1] · Luigi Montrucchio[2] · Giovanni Pistone[3]

Received: 29 May 2018 / Revised: 9 October 2018 / Published online: 20 November 2018

Abstract

The Wasserstein distance on multivariate non-degenerate Gaussian densities is a Riemannian distance. After reviewing the properties of the distance and the metric geodesic, we present an explicit form of the Riemannian metrics on positive-definite matrices and compute its tensor form with respect to the trace inner product. The tensor is a matrix which is the solution to a Lyapunov equation. We compute the explicit formula for the Riemannian exponential, the normal coordinates charts and the Riemannian gradient. Finally, the Levi-Civita covariant derivative is computed in matrix form together with the differential equation for the parallel transport. While all computations are given in matrix form, nonetheless we discuss also the use of a special moving frame.

Keywords Information geometry · Gaussian distribution · Wasserstein distance · Riemannian metrics · Natural gradient · Riemannian exponential · Normal coordinates · Levi-Civita covariant derivative · Optimization on positive-definite symmetric matrices



Volume 1
Issue 2
Article 5

RESEARCH PAPER

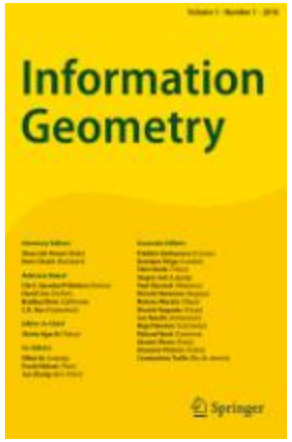
Natural gradient via optimal transport

Wuchen Li[1] · Guido Montúfar[2,3]

Received: 15 March 2018 / Revised: 27 August 2018 / Published online: 19 November 2018

Abstract
We study a natural Wasserstein gradient flow on manifolds of probability distributions with discrete sample spaces. We derive the Riemannian structure for the probability simplex from the dynamical formulation of the Wasserstein distance on a weighted graph. We pull back the geometric structure to the parameter space of any given probability model, which allows us to define a natural gradient flow there. In contrast to the natural Fisher–Rao gradient, the natural Wasserstein gradient incorporates a ground metric on sample space. We illustrate the analysis of elementary exponential family examples and demonstrate an application of the Wasserstein natural gradient to maximum likelihood estimation.

Keywords Optimal transport · Information geometry · Wasserstein statistical manifold · Displacement convexity · Machine learning

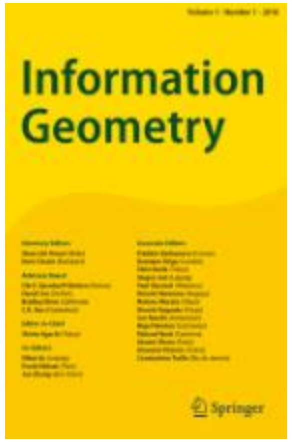


Volume 1
Issue 2
Article 6

RESEARCH PAPER

Superharmonic priors for autoregressive models

Fuyuhiko Tanaka[1]



Volume 1
Issue 2
Article 7

Received: 24 July 2017 / Revised: 9 October 2017 / Published online: 1 December 2017

Abstract Tanaka and Komaki (Sankhya Ser A Indian Stat Inst 73-A:162–184, 2011) proposed superharmonic priors in Bayesian time series analysis as alternative to the famous Jeffreys prior. By definition the existence of superharmonic priors on a specific time series model with finite-dimensional parameter is equivalent to that of positive nonconstant superharmonic functions on the corresponding Riemannian manifold endowed with the Fisher metric. In the autoregressive models, whose Fisher metric and its inverse have quite messy forms, we obtain superharmonic priors in an explicit manner. To derive this result, we developed a systematic way of dealing with symmetric polynomials, which are related to Schur functions.

Keywords Jeffreys prior · Superharmonic priors · Autoregressive models · Noninformative priors · Kullback–Leibler divergence · Fisher metric

RESEARCH PAPER

CrossMark

Asymptotic dependency structure of multiple signals

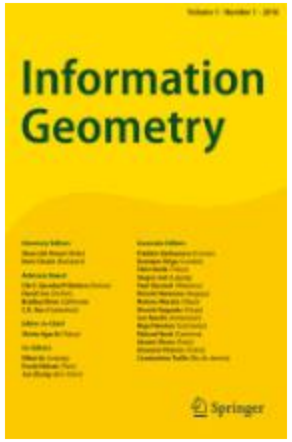
Asymptotic equipartition property for diagrams of probability spaces

Rostislav Matveev[1] · Jacobus W. Portegies[2]

Received: 7 February 2018 / Revised: 16 August 2018 / Published online: 12 October 2018

Abstract

We formalize the notion of the dependency structure of a collection of *multiple* signals, relevant from the perspective of information theory, artificial intelligence, neuroscience, complex systems and other related fields. We model multiple signals by commutative diagrams of probability spaces with measure-preserving maps between some of them. We introduce the asymptotic entropy (pseudo-)distance between diagrams, expressing how much two diagrams differ from an information-processing perspective. If the distance vanishes, we say that two diagrams are asymptotically equivalent. In this context, we prove an asymptotic equipartition property: any sequence of tensor powers of a diagram is asymptotically equivalent to a sequence of homogeneous diagrams. This sequence of homogeneous diagrams expresses the relevant dependency structure.



Volume 1
Issue 2
Article 8

RESEARCH PAPER

Asymptotic dependency structure of multiple signals

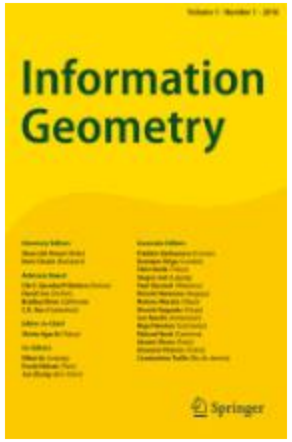
Asymptotic equipartition property for diagrams of probability spaces

Rostislav Matveev[1] · Jacobus W. Portegies[2]

Received: 7 February 2018 / Revised: 16 August 2018 / Published online: 12 October 2018

Abstract
We formalize the notion of the dependency structure of a collection of *multiple* signals, relevant from the perspective of information theory, artificial intelligence, neuroscience, complex systems and other related fields. We model multiple signals by commutative diagrams of probability spaces with measure-preserving maps between some of them. We introduce the asymptotic entropy (pseudo-)distance between diagrams, expressing how much two diagrams differ from an information-processing perspective. If the distance vanishes, we say that two diagrams are asymptotically equivalent. In this context, we prove an asymptotic equipartition property: any sequence of tensor powers of a diagram is asymptotically equivalent to a sequence of homogeneous diagrams. This sequence of homogeneous diagrams expresses the relevant dependency structure.

Keywords Asymptotic equipartition property · Entropy distance · Diagrams of probability spaces · Multiple signals

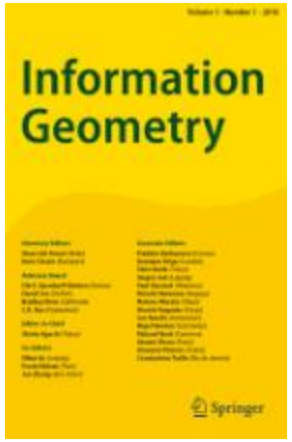


Volume 1
Issue 2
Article 9

RESEARCH PAPER

Ordering positive definite matrices

Cyrus Mostajeran[1] · Rodolphe Sepulchre[1]



Volume 1
Issue 2
Article 10

Received: 3 June 2017 / Revised: 2 May 2018 / Published online: 15 May 2018

Abstract We introduce new partial orders on the set S_n^+ of positive definite matrices of dimension n derived from the affine-invariant geometry of S_n^+. The orders are induced by affine-invariant cone fields, which arise naturally from a local analysis of the orders that are compatible with the homogeneous geometry of S_n^+ defined by the natural transitive action of the general linear group $GL(n)$. We then take a geometric approach to the study of monotone functions on S_n^+ and establish a number of relevant results, including an extension of the well-known Löwner-Heinz theorem derived using differential positivity with respect to affine-invariant cone fields.

Keywords Positive definite matrices · Partial orders · Monotone functions · Monotone flows · Differential positivity · Matrix means