Rho-Tau Embedding of Statistical Models

by Jan Naudts and Jun Zhang

Abstract

This is a review regarding a chapter proposal for publication by Springer in the book Geometric Structures of Information.

Comments to the Authors

The chapter is devoted to the general model of the geometry of rho-tau embedding in statistical models, i.e., statistical manifolds which are smooth and differentiable ones.

The works is remarkably well written and presents an organization that allows the reader to easily go through the basics and understand the rationale of the authors with respect to the existing literature. The work also provides a nice overview in a relevant and important topic on the area of information geometry and divergence models.

The review on the rho-tau divergence is presented in a compact fashion, compatible with the page limit I assume, but allows the reader to get the basics to the following sections. Here, although can be derived, the statement of the generalized Pythagorian equality being satisfied by the rho-tau divergence for any three given points P, Q, R could be linked to a reference.

Also, the definition of the rho-tau entropy in Equation (7) is not well justified. I believe it is a direct extension from the Shannon definition and

usage of general f-divergence, but it would worth mentioning the origin of it so one can easily get the result from Equation (8).

One question that come to my mind is that in such general proposed model one does not need to assume the function (divergence) is Gateaux differentiable in order to ensure the convergence conditions for the expectations?

Regarding the gauge freedom, for the case of $\rho(u) = \frac{1}{\tau'(u)}$ and the corresponding deformed logarithm the divergence becomes a generalization of the Kullback-Leibler one. Would it be possible to select different $\rho(u)$ such that the divergence becomes a general model for the Rényi's divergence? If so, could you comment on that?

With respect to the deformed exponential family (Section 7) if we select a general model, as the one proposed in [17], does the rho-tau divergence can be reduced or extended to the model proposed in [17]? If so, which conditions should be imposed over the deformed family for such situation?