# Differential Geometry for Statistical and Entropy-Based Inference

**Jun Zhang**

**Department of Psychology and Department of Mathematics,**

**The University of Michigan Ann Arbor, USA**

**Email: junz@umich.edu**

Information Geometry is the differential geometric study of the manifold of probability models, and promises to be a unifying geometric framework for investigating statistical inference, information theory, machine learning, etc. Instead of using metric for measuring distances on such manifolds, these applications often use "divergence functions" for measuring proximity of two points (that do not impose symmetry and triangular inequality), for instance Kullback-Leibler divergence, Bregman divergence, f-divergence, etc. Divergence functions are tied to generalized entropy (for instance, Tsallis entropy, Renyi entropy, phi-entropy, U-entropy) and corresponding cross-entropy functions. It turns out that divergence functions enjoy pleasant geometric properties – they induce what is called "statistical structure" on a manifold M: a Riemannian metric g together with a pair of torsion-free affine connections D, D*, such that D and D* are both Codazzi coupled to g while being conjugate to each other. We use these concepts to investigate a generalization of Maximum Entropy principle through conjugate rho-tau embedding mechanism. We show how this generalization captures the various generalization of MaxEnt, including deform-logarithm model and U-model. (Work in collaboration with Jan Naudts [1,2]).

References:

[1] Zhang, J. and Naudts, J. (2017). Information geometry under monotone embedding.Part I: divergence functions. In Nielson, F. and Barbaresco, F. (Eds), *Geometric Science of Information, GSI2017*. *LNCS 10589,* Springer (pp. 205-214).

[2] Naudts, J. and Zhang, J. (2017). Information geometry under monotone embedding.Part II: geometry. In Nielson, F. and Barbaresco, F. (Eds), *Geometric Science of Information, GSI2017*. *LNCS 10589,* Springer (pp. 215-222).