Transfer Learning of Statistical Models on Riemannian Manifolds

T.T. Tran
Université Clermont Auvergne

Supervisor(s): C. Samir

PhD expected duration: Oct. 2020 - Sep. 2023

Address: F109 - ISIMA

E-mail: *tien_tam.tran@doctorant.uca.fr*

Abstract:

Statistical analysis of probability measures is growing in importance in both applications and theory. For instance, for many applications where artificial intelligence is involved, we can deal with data on the space of probability measures. Similarly to standard model, one has to learn from a subset of from training data, and then apply for prediction, classification, segmentation, and retrieval [4]. However, extending standard models from Euclidean spaces to manifolds requires a specific geometric structure and imposes additional constraints. In Information Geometry [2], the space of probability measures, denoted $\mathcal{P}$, is imposed with a metric g and a pair of conjugate connections. As a new result, we show that traditional connection coincides with the Levi-Civita connection.

In this work, we consider the space $\mathcal{P}$ in the finite case with dimension $N+1$. Hence, the space can be represented as the simplex in the Euclidean space $\mathbb{R}^{N+1}$. We introduce the manifold structure on $\mathcal{P}$ by the chart with the usual projection on the first N coordinates. Then we can define the bases for tangent and cotangent spaces. In fact, one remarkable property is that these spaces are trivial bundles. In the literature [3], give the definition for the statistical bundle with a basis of $\mathcal{P}$. The Exponential atlas and the Mixture atlas were defined on the statistical bundle, so that the tangent space is the corresponding fiber of the statistical bundle. Unfortunately, this statistical bundle is not trivial.

Many different metrics, semi-metrics on $\mathcal{P}$ have been studied, including Wassertein, Jensen-Shannon, $\mathbb{L}^1$, $\mathbb{L}^2$. In this work, the space $\mathcal{P}$ is then equipped with the Fisher-Rao metric g^{FR} and the the Levi-Civita connection. Using a local coordinate system, we introduce the Christoffel symbols, the geodesic, the exponential map, the logarithmic map and the parallel transport. We note that it is not straightforward to solve for the geodesic and parallel transport in a general Riemannian manifold. Furthermore, the Riemannian space $(\mathcal{P}, g^{FR})$ is proved to be isometric with the unit Sphere in $\mathbb{R}^{N+1}$ which allow us to use nice and practical properties.

As an application, we focus on the problem of transfer learning [1]. Usually, data are mapped to the tangent space of a specific position. The idea of transfer learning is to transfer the statistical model from an original point to another point, called target point. In particular, we propose a transfer learning using explicit expressions of parallel transport. We show different applications where transferring a learned statistical model is of interest: Principle Component Analysis (PCA), and Linear Regression. The results show the effectiveness of the proposed methods with potential extension to different real world problems.

References

[1] O. Freifeld, S. Hauberg, and M. J. Black. Model transport: Towards scalable transfer learning on manifolds. *CVPR*, pages 1378–1385, 2014.

[2] Frank Nielsen. An elementary introduction to information geometry. *Entropy*, 22(10), 2020.

[3] G. Pistone. Information geometry of the probability simplex: A short course. *Nonlinear Phenomena in Complex Systems*, pages 221–242, 2020.

[4] A. Srivastava, I. Jermyn, and S. Joshi. Riemannian analysis of probability density functions with applications in vision. *CVPR*, 2007.

Short biography – We are interested in deploying Gaussian Processes in Riemannian manifolds, such as the space of probability measures. To do so, we need to study the Riemannian structure and Covariance functions on this space. Then we want to apply it in machine vision to Detecting and Classifying multiple forms of objects. We aim to develop accurate methods that will be able to learn and do tasks in a single unified framework. This PhD thesis is funded by ANR-AI program of Clermont Auvergne University.