Abstract

Manifold-valued data naturally arises in medical imaging. In cognitive neuroscience, for instance, brain connectomics bases the analysis of coactivation patterns between different brain regions on the analysis of the correlations of their functional Magnetic Resonance Imaging (fMRI) time series – an object thus constrained by construction to belong to the manifold of symmetric positive definite matrices.

One of the challenges that naturally arises in these studies consists of finding a lower-dimensional subspace for representing such manifold-valued and typically high-dimensional data. Traditional techniques, like principal component analysis, are ill-adapted to tackle non-Euclidean spaces and may fail to achieve a lower-dimensional representation of the data – thus potentially pointing to the absence of lower-dimensional representation of the data. However, these techniques are restricted in that: (i) they do not leverage the assumption that the connectomes belong on a pre-specified manifold, therefore discarding information; (ii) they can only fit a linear subspace to the data.

For this talk, we are interested in learning potentially highly curved submanifolds of manifold-valued data. Motivated by the brain connectomes example, we investigate a latent variable generative model, which has the added benefit of providing us with uncertainty estimates – a crucial quantity in the medical applications we are considering. While latent variable models have been proposed to learn linear and nonlinear spaces for Euclidean data, or geodesic subspaces for manifold data, fewer intrinsic latent variable models exist to learn nongeodesic subspaces for manifold data. We formulate a Riemannian variational autoencoder with an intrinsic generative model of manifold-valued data. We evaluate its performances on synthetic and real datasets by introducing the formalism of weighted Riemannian submanifolds.