# Xavier Pennec

## Barycentric Subspace Analysis: an extension of PCA to Manifolds

I address in this talk the generalization of Principal Component Analysis (PCA) to Riemannian manifolds and potentially more general stratified spaces. Tangent PCA is often sufficient for analyzing data which are sufficiently centered around a central value (unimodal or Gaussian-like data), but fails for multimodal or large support distributions (e.g. uniform on close compact subspaces). Instead of a covariance matrix analysis, Principal Geodesic Analysis (PGA) and Geodesic PCA (GPCA) are proposing to minimize the distance to Geodesic Subspaces (GS) which are spanned by the geodesics going through a point with tangent vector is a restricted linear sub-space of the tangent space. Other methods like Principal Nested Spheres (PNS) restrict to simpler manifolds but emphasize on the need for the nestedness of the resulting principal subspaces.

In this work, we first propose a new and more general type of family of subspaces in manifolds that we call barycentric subspaces. They are implicitly defined as the locus of points which are weighted means of $k+1$ reference points. As this definition relies on points and do not on tangent vectors, it can also be extended to geodesic spaces which are not Riemannian. For instance, in stratified spaces, it naturally allows to have principal subspaces that span over several strata, which is not the case with PGA. We show that barycentric subspaces locally define a submanifold of dimension $k$ which generalizes geodesic subspaces. Like PGA, barycentric subspaces can naturally be nested, which allow the construction of inductive forward nested subspaces approximating data points which contains the Fréchet mean. However, it also allows the construction of backward flags which may not contain the mean. Second, we rephrase PCA in Euclidean spaces as an optimization on flags of linear subspaces (a hierarchies of properly embedded linear subspaces of increasing dimension). We propose for that an extension of the unexplained variance criterion that generalizes nicely to flags of barycentric subspaces in Riemannian manifolds. This results into a particularly appealing generalization of PCA on manifolds, that we call Barycentric Subspaces Analysis (BSA).