

# Principal Boundary on Riemannian Manifolds and Classification Problem

Zhigang Yao

*National University of Singapore*

*E-mail: zhigang.yao@nus.edu.sg*

**Abstract:** We will discuss the problem of finding principal components to the multivariate datasets, that lie on an embedded nonlinear Riemannian manifold within the higher-dimensional space. Our aim is to extend the geometric interpretation of PCA, while being able to capture the non-geodesic form of variation in the data. We introduce the concept of a principal sub-manifold, a manifold passing through the center of the data, and at any point of the manifold, it moves in the direction of the highest curvature in the space spanned by the eigenvectors of the local tangent space PCA. We show the principal sub-manifold yields the usual principal components in Euclidean space. We illustrate how to find, use and interpret the principal sub-manifold, with which a classification boundary can be defined for data sets on manifolds.