

A theory of stratification learning : Dimension-wise clustering with reconstruction

Abstract : Given i.i.d. sample from a stratified mixture of immersed manifolds of different dimensions, we will study the minimax estimation of the underlying stratified structure. We will provide a constructive algorithm allowing to estimate each mixture component at its optimal dimension-specific rate adaptively. The method is based on an ascending hierarchical co-detection of points belonging to different layers, which also identifies the number of layers and their dimensions, assigns each data point to a layer accurately, and estimates tangent spaces optimally. These results hold regardless of any ambient assumption on the manifolds or on their intersection configurations. They open the way to a broad clustering framework, where each mixture component models a cluster emanating from a specific nonlinear correlation phenomenon.