

1164-52-176

Braxton Osting*, Department of Mathematics, University of Utah, 155 S 1400 E RM 233, Salt Lake City, UT 84112. *Consistency of archetypal analysis.*

Archetypal analysis is an unsupervised learning method that uses a convex polytope to summarize multivariate data. For fixed k , the method finds a convex polytope with k vertices, called archetype points, such that the polytope is contained in the convex hull of the data and the mean squared distance between the data and the polytope is minimal. In this talk, I'll give a consistency result that shows if the data is independently sampled from a probability measure with bounded support, then the archetype points converge to a solution of the continuum version of the problem, of which we identify and establish several properties. We also obtain the convergence rate of the optimal objective values under appropriate assumptions on the distribution. If the data is independently sampled from a distribution with unbounded support, we also prove a consistency result for a modified method that penalizes the dispersion of the archetype points. Our analysis is supported by detailed computational experiments of the archetype points for data sampled from the uniform distribution in a disk, the normal distribution, an annular distribution, and a Gaussian mixture model. This is joint work with Dong Wang, Yiming Xu, and Dominique Zosso. (Received January 18, 2021)