T10c Galaxy Group Finding via Unsupervised Machine Learning

Hai-Xia MA (1), Tsutomu T. TAKEUCHI (1, 2), Suchetha COORAY (1)
(1) Nagoya University; (2) Institute for Statistical Mathematics

As one of the most popular unsupervised classification techniques, clustering is the task of dividing the population or data points into a number of groups such that data points in the same groups are more similar to other data points in the same group than those in other groups. In an astrophysical context, clustering is to assign galaxies into galaxy groups. In this work, We create a sample of identified galaxy groups with $0.01 < z < 0.20$ from the Sloan Digital Sky Survey (SDSS) Data Release 7 (DR7). We apply a variety of unsupervised machine learning methods: Dynamic Friends-of Friends (DFoF), Gaussian Mixture Models and its variants (GMMs), and Ordering points to identify the clustering structure (OPTICS). We perform a test of the clustering results obtained by different methods in terms of completeness and contamination. With fixed free parameters, all the clustering algorithms produce the best group catalogues as quantified by the statistics. The results demonstrate that the OPTICS and Extreme Deconvolution GMM (XDGMM) are robust group finders that are efficient in determining a wide variety of group shapes and sizes with very low contamination. We also investigate the gravitational dynamics and group properties of these galaxy groups.