

Statistics Seminar
Department of Mathematical Sciences

DATE:	Friday, April 24, 2015 (NOTE THE UNUSUAL DAY)
TIME:	2:10pm—3:10pm (NOTE THE UNUSUAL TIME)
LOCATION:	WH 100E
SPEAKER:	Yifan Xu (Case Western Reserve University)
TITLE:	Fast clustering using adaptive density peak detection

Abstract

Common limitations of clustering methods in bioinformatics include the slow computation and algorithm convergence, the prespecification of a number of different intrinsic parameters, and the lack of robustness. Rodriguez and Laio ["Clustering by fast search and find of density peaks", *Science*, vol. 344, no. 6191 pp. 1492-1496, June 2014] proposed a fast clustering algorithm, which is based on search of cluster centers through their local densities. However, the selection of the key intrinsic parameters in the algorithm was not systematically investigated. It is relatively difficult to estimate the "optimal" parameters since the original definition of the local density in the algorithm is based on a truncated counting measure.

We propose a clustering procedure with adaptive density peak detection, where the local density is estimated through the nonparametric multivariate kernel estimation. We also develop an automatic cluster centroid selection method through maximizing an average silhouette index. The advantage and flexibility of the proposed method are demonstrated through simulation studies and the analysis of a few benchmark gene expression data sets. The method only needs to perform in one single step without iteration, and can be parallelized, having a great potential to apply on large data sets. We also made an R-package "ADPclust" that is available on CRAN soon.

This is a joint work with Xiao-Feng Wang from Department of Quantitative Health Sciences at Cleveland Clinic Foundation.

From:

<http://www2.math.binghamton.edu/> - **Department of Mathematics and Statistics, Binghamton University**

Permanent link:

<http://www2.math.binghamton.edu/p/seminars/stat/04242015>



Last update: **2015/04/20 14:55**