



THE UNIVERSITY OF  
CHICAGO

Department of Statistics

MASTER'S THESIS PRESENTATION

---

ZHENGJIAN SONG

Department of Statistics  
The University of Chicago

Clusterpath: A Convex Clustering Method

THURSDAY, May 15, 2014, at 8:30 AM  
Eckhart 117, 5734 S. University Avenue

#### ABSTRACT

The recently introduced Clusterpath is a new method for clustering. As a convex relaxation of  $k$  means and agglomerative hierarchical clustering, it casts the clustering task to a family of convex optimization problems. Compared to standard methods such as  $k$ -means, hierarchical clustering, Gaussian mixture models and spectral clustering, it provides a more natural geometric interpretation of the clustering behaviors through a tree structure learnt from the data. In this paper, we implement two efficient algorithms, the Subgradient decent algorithm and the Alternating Minimization Algorithm (AMA), for calculating the continuous regularization path of solutions. By simulation, we verify that the AMA outperforms the Subgradient decent algorithm in solving the convex clustering problem. Our empirical experiments also show that the Clusterpath can give state-of-art results similar to spectral clustering. Moreover, we propose a new heuristic method for choosing the number of clusters automatically within the Clusterpath framework using a pattern discovered in the solution path.

---

For information about building access for persons with disabilities, please contact Matt Johnston at 773.702-0541 or send an email to [mhj@galton.uchicago.edu](mailto:mhj@galton.uchicago.edu).

If you wish to subscribe to our email list, please visit the following web site:  
<https://lists.uchicago.edu/web/arc/statseminars>.