Convex Co-clustering of Tensors

Eric Chi, PhD
Assistant Professor
Department of Statistics
North Carolina State University

ABSTRACT

Clustering is a fundamental unsupervised learning technique that aims to discover groups of objects in a dataset. Biclustering extends clustering to two dimensions where both observations and variables are grouped simultaneously, such as clustering both cancerous tumors and genes or both documents and words. In this work, we develop and study a convex formulation of the generalization of biclustering to co-clustering the modes of multiway arrays or tensors, the generalization of matrices. Our convex co-clustering (CoCo) estimator is guaranteed to obtain a unique global minimum of the formulation and generates an entire solution path of possible co-clusters governed by a single tuning parameter. We extensively study our method in several simulated settings, and also apply it to an online advertising dataset. We also provide a finite sample bound for the prediction error of our CoCo estimator.

This is joint work with Brian Gaines, Wei Sun, and Hua Zhou.

12:00 Noon, **Tuesday, October 3, 2017**
LEPH 115, 60 College Street
Lunch at 11:45 AM in LEPH 108