



THE UNIVERSITY OF
CHICAGO

Department of Statistics

MASTER'S THESIS PRESENTATION

DONGYUE XIE

Department of Statistics
The University of Chicago

Generalized Sparse Canonical Correlation Analysis with
Applications to Microbiome Data

TUESDAY, February 6, 2018, at 9:00 AM
Jones 304, 5747 S. Ellis Avenue

ABSTRACT

Human microbiome is all the microbes in and on human body and has huge impact on health and disease. High-throughput sequencing technologies have enabled researchers to study the microbiome composition. Motivated by studying the relationships between two or more microbiome compositions and improving interpretability of results, we proposed a new method for testing and finding the association relationships between two random vectors in a high-dimensional setting. The new method is based on sparse canonical correlation analysis(SCCA) and has two special features. Firstly, it could take the general structure of variables into account. Phylogenetic tree is important prior knowledge of evolutionary relationships among bacterial taxa and our method could incorporate tree structured penalty in SCCA. Secondly, it can be easily transformed into a generalized lasso problem and solved efficiently. The proposed method was applied to the newest microbiome data from American Gut Project.

For information about building access for persons with disabilities, please contact Laura Rigazzi at 773.702-0541 or send an email to lrigazzi@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>.