

Statistics Seminar
Department of Mathematical Sciences

DATE:	Thursday, Nov. 12, 2020
TIME:	1:15pm - 2:15pm
LOCATION:	Zoom meeting
SPEAKER:	Xiaoke Qin, Binghamton University
TITLE:	Variable selection for sparse Dirichlet-Multinomial regression with an application to microbiome data analysis.

Abstract

With the development of next generation sequencing technology, researchers have now been able to study the microbiome composition using direct sequencing, whose output are bacterial taxa counts for each microbiome sample. One goal of microbiome study is to associate the microbiome composition with environmental covariates. This paper proposes to model the taxa counts using a Dirichlet-multinomial (DM) regression model in order to account for overdispersion of observed counts. The DM regression model can be used for testing the association between taxa composition and covariates using the likelihood ratio test. However, when the number of covariates is large, multiple testing can lead to loss of power. To address the high dimensionality of the problem, a penalized likelihood approach is proposed to estimate the regression parameters and to select the variables by imposing a sparse group l_2 penalty to encourage both group-level and within-group sparsity. Such a variable selection procedure can lead to selection of the relevant covariates and their associated bacterial taxa. An efficient block-coordinate descent algorithm is developed to solve the optimization problem in this paper. The authors also demonstrate the power of the method in the analysis of a data set evaluating the nutrient intake on the human gut microbiome.

From:

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

Permanent link:

<https://www2.math.binghamton.edu/p/seminars/stat/201112>

Last update: **2020/11/03 19:44**

