MICHIGAN STATE UNIVERSITY
Department of Statistics and Probability

COLLOQUIUM

Yoonkyung Lee

Ohio State University

Generalized Principal Component Analysis: Dimensionality Reduction through the Projection of Natural Parameters

Tuesday, November 10, 2015
10:20 a.m. - 11:10 am
Refreshments 10:00 am
C405 Wells Hall

Abstract

Principal component analysis (PCA) is useful for a wide range of data analysis tasks. However, its implicit link to the Gaussian distribution can be undesirable for discrete data such as binary and multi-category responses or counts. We generalize PCA to handle various types of data using the generalized linear model framework. In contrast to the existing approach of matrix factorizations for exponential family data, our generalized PCA provides low-rank estimates of the natural parameters by projecting the saturated model parameters. Due to this difference, the number of parameters does not grow with the number of observations and the principal component scores on new data can be computed with simple matrix multiplication. We provide a computationally efficient algorithm for finding the principal component loadings and demonstrate the benefits of the proposed approach numerically.

This is joint work with Andrew Landgraf.

To request an interpreter or other accommodations for people with disabilities, please call the Department of Statistics and Probability at 517-355-9589.