The University of Chicago
Department of Statistics

MASTER'S THESIS PRESENTATION

XIAOCHUN NI

Department of Statistics
The University of Chicago

Curve Clustering of Time Course Gene Expression Data Using Gaussian Process

WEDNESDAY, February 23, 2011, at 2:00 PM

110 Eckhart Hall, 5734 S. University Avenue

ABSTRACT

Time course gene expression data comprising transcript abundance measurements of large number of genes at various time points is a typical type of data generated in the files of biology and medical research. One major goal for such study is to find clusters of genes with similar temporal expression pattern. The current commonly used methods usually oversimplify the data structure. In this paper, we model the data using Gaussian Process with a spline mean and propose two clustering procedures. Using a microarray expression dataset of 30 genes along fruit fly embryogenesis (31 time points), we apply previous methods and our own procedures and discuss the result.

Information about building access for persons with disabilities may be obtained in advance by calling Sandra Romero at 773.702-0541 or by email (sandra@galton.uchicago.edu).