The University of Chicago
Department of Statistics

MASTER'S THESIS PRESENTATION

LAN WEI
Department of Statistics
The University of Chicago

Maximum-Likelihood Estimate for Distance Matrix

WEDNESDAY, July 6, 2011 at 11:00 AM
110 Eckhart Hall, 5734 S. University Avenue

ABSTRACT

The structure of relationship among subjects is an interesting topic in many social studies, especially phylogeny. The between-subject relationships are commonly measured by similarity or dissimilarity metrics and represented using a distance matrix. Based on this distance matrix, a tree can be constructed to characterize the overall structure of inferred relationships among all subjects. Currently, there are various projection methods that can estimate a tree from an observed distance matrix. However, there exists no broad agreement about objective criteria to quantify the goodness-of-fit of the inferred tree.

In this study, we suggest a new method that improves current tree projection algorithms by finding the global maximum of the likelihood in the whole projection space. Our method is based on a log likelihood function for distance matrices in a closed form where the Newton-Raphson algorithm can be directly applied. By perturbing the observed distance matrices, our procedure computes the deviance of the fitted tree and identifies the optimal tree. Empirical studies were carried out using two widely used tree projection algorithms, namely Neighbor-Joining (NJ) and Unweighted Pair Group Method with Arithmetic Mean (UPGMA). Our results suggest that the proposed method can improve the estimate of the tree.

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).