

Differential Expression Analysis Using Hierarchical Mixture Models

Michael A. Newton
University of Wisconsin, Madison

Mixture modelling provides an effective approach to the differential expression problem in microarray data analysis. Methods based on fully parametric mixture models are available, but lack of fit in some examples indicates that more flexible models may be beneficial. Existing, more flexible, mixture models work at the level of one-dimensional gene-specific summary statistics, and so when there are relatively few measurements per gene these methods may not provide sensitive detectors of differential expression. I describe a semiparametric hierarchical mixture model to provide methodology that is both sensitive in detecting differential expression and sufficiently flexible to account for the complex variability of normalized microarray data. I consider the two sample comparison problem in detail; an experiment involving Affymetrix GeneChips and yeast translation provides the motivating case study. A recursive computational method is described to fit the semiparametric model. Gene-specific posterior probabilities of differential expression form the basis of statistical inference; I show how they define short gene lists and also how they provide estimates of the rate of false discoveries. Compared to both gene-specific t-tests and the method of Efron et al. (2001), the proposed methodology exhibits better operating characteristics in a simulation study.

Details of a parametric submodel are also provided.

Reference

This talk is based on Technical Report #1074, Department of Statistics, University of Wisconsin, Madison, "Detecting differential gene expression with a semiparametric hierarchical mixture method." by M.A. Newton, A. Nouiery, D. Sarkar, and P. Ahlquist.
<http://www.stat.wisc.edu/~newton/research/arrays.html>