

CLUSTER ANALYSIS OF TIME-COURSE MICROARRAY DATA USING TRANSFORMATION

Mira Park¹, Hyo-Jung Lee², Peol-A Kim²

¹*Department of Preventive Medicine, Eulji University, KOREA*

²*Department of Statistics, Korea University, KOREA*

One of the main purposes of the microarray analysis is to find clusters of genes or samples with similar expression patterns. To analyze time-course microarray data, several clustering methods such as hierarchical clustering, k-means clustering and SOM have been applied to the raw expression data. However, these approaches have several problems because they treat each time points as independent variables. They ignore the design of study, as no account is taken of the fact that the values at each time point are from same individual. And also they do not consider orders and intervals between time points. To avoid this problem, we benchmark the simple idea for analyzing pharmacokinetics data. We transform the data before clustering using summary statistics: height of peak, peak time, and area under curve (AUC). We demonstrate the usability of the method by real microarray data and simulated serial data. We also compare the results with raw data analysis and discuss the merits and demerits of the proposed method.