

Multidimensional scaling for large genomic data sets

Henry Horng-Shing Lu

Institute of Statistics, National Chiao Tung University, Hsinchu, Taiwan.

Abstract

We developed a new rapid metric multidimensional scaling (MDS) method with a low computational complexity. Computer simulation showed that the new method of split-and-combine MDS (SC-MDS) is fast, accurate and efficient. In the empirical studies evaluated in this report using microarray data on the yeast cell cycle, we found that the performance of K-means in the reduced dimensional space is better than that of K-means in the original space. When we remove the back-ground noise by projecting expression profiles to the reduced dimensional MDS space, the clustering result reveals clearly the pattern of two cell cycles. The biological validation also confirms the advantage of SC-MDS. Hence, the proposed SC-MDS is useful for biological investigation using whole genome data. This is a joint work with Dr. Jengnan Tzeng and Wen-Hsiung Li.