

FIMXII-SCMA2005@AUBURN, Twelfth Annual International Conference on Statistics, Combinatorics, Mathematics and Applications, December 2–4, 2005, Auburn University, Auburn, Alabama, USA

Coauthors: Nader Ebrahimi, Department of Statistics, Northern Illinois University

**MOLECULAR PATTERN DISCOVERY USING NON-NEGATIVE
MATRIX FACTORIZATION BASED ON RENYI'S
INFORMATION MEASURE**

KARTHIK DEVARAJAN

Current efforts in cancer research focus on predicting drug responses of patients by analyzing gene expression patterns and identifying genes that cause specific responses, with a view to personalizing cancer treatment. DNA microarray technology has made it possible to simultaneously measure the expression levels of tens of thousands of genes. A related problem is that of class discovery, which involves unsupervised clustering of tissue samples or genes based on corresponding gene expression data and to identify the genes that distinguish the clusters. Non-negative matrix factorization (NMF) is a powerful method for decomposing the gene expression matrix V into two matrices with positive entries, $V = WH$, where each column of W defines a metagene and each column of H represents the metagene expression pattern of the corresponding sample (Brunet et al, Bioinformatics, 2004). We propose an unsupervised method for molecular pattern discovery using NMF, based on Renyi's measure of distance between two non-negative matrices, related to the Poisson likelihood. This is a generalization of previous work (Lee Seung, Nature, 1999), and includes various well-known distance measures as special cases. We demonstrate the applicability of this method using cancer microarray data.

ASSOCIATE MEMBER, DEPARTMENT OF BIostatISTICS, FOX CHASE CANCER CENTER, PHILADELPHIA

E-mail address: `karthik.devarajan@fccc.edu`