STANFORD UNIVERSITY
DEPARTMENT OF STATISTICS

DEPARTMENTAL SEMINAR

4:15 p.m., Tuesday, March 4, 2008
Sequoia Hall Room 200
(Cookies at 3:45 in 1st Floor Lounge)

Peter Hoff
Departments of Statistics and Biostatistics
University of Washington

Hierarchical eigenmodels for matrix-valued data

Matrix decomposition models are a popular way to represent matrix-valued data. For example, the variation among the entries of an $n \times p$ data matrix Y is often expressed using a singular value decomposition model $Y \sim UDV^T + E$, where U and V are orthonormal matrices and D is a diagonal matrix. In this work I consider pooling information across multiple such data matrices $Y^{(1)}, \ldots Y^{(K)}$ for situations in which common columns across matrices represent repeated measurements under a common set of conditions. For example, suppose individuals from several different groups have expression levels measured across a common set of genes. If $y_{i,j}^{(k)}$ is the expression level of gene j for patient i in group k, then it might be beneficial to share information about the column variation across the different groups. I propose doing this by estimating the parameters in a model for the variability among the orthonormal eigenvector matrices $V^{(1)}, \ldots, V^{(K)}$ of the K data matrices. The model is based on a variation of the matrix Bingham distribution, for which estimation is accomplished primarily with Gibbs sampling. The methodology is applied to the analysis of several multiple multivariate datasets.