

STANFORD UNIVERSITY
DEPARTMENT OF STATISTICS
DEPARTMENTAL SEMINAR

4:15 p.m., Wednesday, February 14, 2001
Sequoia Hall Rm. 200

Allan Aasbjerg Nielsen
Informatics and Mathematical Modelling
Technical University of Denmark, Building 321
DK-2800 Kongens Lyngby, Denmark

Multi-Set Canonical Correlations Analysis

This talk deals with extensions of the established two-set canonical correlations analysis. In two-set analysis data naturally divides into two multivariate groups. Based on the original zero-mean variables orthogonal linear combinations with maximal correlations, the so-called canonical variates, are found. In multi-set analysis the maximization of correlation, which is a scalar, to find two sets of canonical variates is replaced by optimization of different measures of the correlation matrix of more sets of canonical variates. This optimization may be maximization of the sum of correlations, maximization of the sum of squared correlations, maximization of the largest eigenvalue, minimization of the smallest eigenvalue, or minimization of the generalized variance.

Maximization of the sum of correlations will be used to give two examples of remote sensing applications, one terrestrial and one oceanic. The terrestrial case will use annual Landsat TM data from 1984 to 1989 to look into change in a forested region in northern Sweden. This case will also give a brief comparison between the five optimizations mentioned above. The oceanic case will use monthly means of global sea surface temperatures from the NOAA/NASA AVHRR Oceans Pathfinder data covering 1987 to 1998 to reveal patterns in global temperature variation including El Nio and La Nia.