

Sparse latent factor analysis in post-genomic data fusion

Peter J Green,
University of Bristol, UK

Abstract

Classical latent factor analysis seeks to discover patterns of dependence in multivariate data that allow dimension reduction through the representation of the observed variables as linear combinations of a smaller number of unobserved 'factors'. We are interested in finding sparse representations, in which there are many zero coefficients among the linear coefficients, in the interests of parsimony, interpretability, and statistical stability; we use a Bayesian hierarchical modelling approach. Specifically, we examine the situation where there are two or more groups of variables, neither low in dimension, and the main interest is in discovering sparse representations of the dependence between them. We are motivated by a study relating profiling of metabolites with transcript and enzyme activity.