Working Group on Statistical Learning Seminar

Title: Latent Variable Models for Ordinal Data

Speaker: Damien McParland

Date: Fri 1st April 2011 at 1:00PM

Location: Statistics Seminar Room- L550 Library building

Abstract: Ordinal data arise in many contexts and item response modelling is a long established method for analysing this type of data.

The ordinal response for individual i on item j is denoted Y_{ij}, where $i = 1, \ldots, N$ and $j = 1, \ldots, J$. Corresponding to each ordinal data point Y_{ij} is a latent Gaussian variable Z_{ij}. The value of Y_{ij} is observed to be level k if the latent Gaussian variable Z_{ij} lies within a specified interval. In addition, another latent Gaussian variable θ_i, often called a latent trait, is used to model the underlying attributes of individual i. The mean of Z_{ij} depends on θ_i, i.e.

$$Z_{ij} \sim \mathrm{N}(a_j\theta_i - b_j, 1)$$

In the item response literature, a_j and b_j are typically known as discrimination and difficulty parameters respectively.

The extension to a mixture of two parameter item response models, which provides clustering capabilities in the context of ordinal data is also explored. In this context

the mean of Z_{ij} also depends on which group individual i belongs to, i.e.

$$Z_{ij} \sim \mathsf{N}(a_{gj}\theta_i - b_{gj}, 1)$$

where a_{gj} and b_{gj} are group specific discrimination and difficulty parameters.

Estimation of both of these models within the Bayesian paradigm is achieved using a Metropolis-within-Gibbs sampler.