Bayesian methods
Lab notes

Lab 7 Miscellanea in BUGS: a probit model via latent variable and Stochastic Search Variable Selection via a hierarchical normal mixture model

`Example 1` A probit model via latent variable

`Reference` Example taken from from Finney, D. J. (1947) *The estimation from individual records of the relationship between dose and quantal response*, *Biometrika*, **34**, Issue 3/4, 320-334.

`Language` BUGS code: *Bernoulli data: via probit/logit or via latent variable* in `latentv.b`.
Reference code: *Program 4.12 SATM scores*, example 4.12 in Congdon's book.

`Subject` *Finney's data give the presence or absence of vasoconstriction in the skin of the fingers following inhalation of a certain volume of air at a certain average rate. After investigating the data, the candidate models for analysing the (two dimensional log)dose-quantal response relationship are the usual probit and logit models. Once established the equivalence between the usual probit model and the latent normal variable model for dichotomous response, we compute the full conditionals for the two alternative formulations. Then, we run Gibbs sampling using WinBUGS to compare the various models.*

`Example 2` Stochastic Search Variable Selection via a hierarchical normal mixture model

`Reference` Example 4.1 in George, E. I. and McCulloch R. E. (1993) *Variable Selection Via Gibbs Sampling*, *JASA*, **88**, Issue 423, 881-889.

`Language` BUGS code: *Stochastic Search Variable Selection. Simulated data example* in `SSVSsim.b`.
Reference code: *Program 4.25 Two-stage variable selection with simulated data*, Example 4.25 in Congdon's book, p. 141.

`Subject` *A crucial problem in building a multiple regression model is the selection of predictors to include. That is, given a dependent variable Y and a set of potential predictors $X_1, \ldots, X_p$, the problem is to find the 'best' model of the form $Y = X_1^*\beta_1^* + \ldots + X_q^*\beta_q^* + \epsilon$, where $X_1^*, \ldots, X_q^*$ is a 'selected' subset of $X_1, \ldots, X_p$. There are 2^p potential regression models. A wide variety of selection procedures are based on a comparison of all 2^p possible submodels. Instead, SSVS*

is a procedure to select 'promising' subsets - which may be not necessarily 2^p - for further considerations. It is based on embedding the entire regression setup in a hierarchical Bayes normal mixture model, where latent variables are used to identify subset choices. The Gibbs Sampler is used to update the initial probabilities assigned to the different subset choices. .

The BUGS codes and the datasets are linked at the course web page.