

Fast Bayesian Implementation of Hierarchical Mixtures of Experts and Stochastic Neural Networks

Samuel P. Wong

Assistant Professor

Department of Information & Systems Management

Hong Kong University of Science & Technology

imsam@ust.hk

Keywords: Stochastic Neural Networks, Hierarchical Mixtures of Experts, Bayesian Computation, Gibbs Sampler, Parameters Expansion.

Abstract

Hierarchical Mixtures of Experts (HME) (Jordan & Jacobs 1994) and Stochastic Neural Networks (SNN) (Lai & Wong 2001) are closely related technique in nonparametric function estimation. The maximum likelihood estimates of both models are usually computed via the Expectation–Maximization (EM) algorithm. Despite the successes in many reported cases (Ripley 1996 p. 283-6), the E-step computation is exponential in the number of experts/neurons and therefore, constrain their applications to relatively small data sets and/or simple HME/SNN structures. The Bayesian implementation of HME and SNN reduces the computational burden by using Gibbs Sampler. Employing the newly proposed parameters expansion technique, we significantly speed up the convergence of the Gibbs Sampler. Some simulated and real examples will be presented.

References

- Lai, T.L. & Wong, S.P. (2001) “Stochastic neural networks with applications to nonlinear time series.” *Journal of American Statistical Association*, **96**, 968-981.
- Jordan, M.I., Jacobs, R.A. (1994) “Hierarchical mixtures of experts and the EM algorithm”, *Neural Computation*, **6**, 181-214.
- Ripley, B.D. (1996) *Pattern Recognition and Neural Networks*. Cambridge University Press.