Cross-Validation, Information Criteria, Expected Utilities and the Effective Number of Parameters

Aki Vehtari and Jouko Lampinen
Laboratory of Computational Engineering
Helsinki University of Technology
P.O.Box 9400, FIN-02015, HUT, Finland
{Aki.Vehtari,Jouko.Lampinen}@hut.fi

Abstract. We consider practical methods for Bayesian model assessment and selection based on expected utilities. We discuss relations between cross-validation and information criteria, which can be used to estimate the expected utility of the model. Specific criteria discussed are Akaike's (AIC), Bayesian (BIC), network (NIC), and deviance information criteria (DIC). We also discuss the concept of the effective number of parameters, how it is estimated in information criteria and describe how it can be estimated using the cross-validation approach.

Cross-validation approach approximates the expected utility using the cross-validation predictive densities, while information criteria approximate the expected utility asymptotically, which will not necessarily give good approximation in a case of complex hierarchical model and finite data. Cross-validation approach uses full predictive distribution obtained by integrating out the unknown parameters, while information criteria use a *plug-in* predictive distributions (maximum likelihood, maximum a posteriori or posterior mean), which ignore the uncertainty about parameter values and model. In cross-validation approach the distribution of the estimate is relatively easy to obtain, while it is not so easy for information criteria and they are usually used to give only a point estimate of the expected utility, which leads to selection of unnecessarily large models.

Essential part of the modern information criteria (like DIC) is the estimation of the effective number of parameters p_{eff} in model. Estimation of the p_{eff} is especially difficult for complex hierarchical models. In the cross-validation approach, the estimate of the p_{eff} is not needed for model assessment or comparison, but it may provide additional insights to models. We show that p_{eff} can be easily estimated by the difference of the expected posterior predictive likelihood and the expected predictive likelihood estimated with the cross-validation approach.