On Bayesian uncertainty quantification

Aad van der Vaart[1]

[1] *Universiteit Leiden, Netherlands, avdvaart@math.leidenuniv.nl*

Keywords: Credible set, confidence set, nonparametric estimation, Bayesian adaptation, sparsity

We give an overview of theoretical results that justify, or not, the use of a posterior distribution of a high- or infinite-dimensional parameter as a method for uncertainty quantification. The posterior distribution is the conditional distribution of the parameter given the data when the parameter is thought of as having been generated by a prior. For the non-Bayesian it is just a random distribution over the parameter space, whose spread could be used to form the equivalent of a confidence set. We start by noting that in the nonparametric situation this is only justified if the prior does not oversmooth the true parameter. In practice, the smoothness of the prior is chosen dependent on the data so as to adapt the procedure to unknown smoothness or sparsity level. We show that then uncertainty quantification is correct for a large set of, but not all, true parameters. This is related to the incompatibility of adaptation and uniform coverage by any statistical procedure, including Bayesian ones. We illustrate the results by the examples of Gaussian process priors for curve or surface estimation and inverse problems, Dirichlet mixtures for density estimation, and the spike-and-slab and horseshoe priors for the estimation of high-dimensional sparse models, every of which we briefly introduce.