Luca Tardella

Some Topics In Bayesian Methodology

Michael Lavine, Advisor

ABSTRACT

The thesis is concerned with the problem of extending the methodology for robustness investigations in the framework of nonparametric Bayesian inference. Robustness here is intended in the sense of the determination of the global variation of outputs given an imprecise formulation of the inputs. Considering nonparametric models and robustness jointly allows an escape from the possible -- inevitable -- sources of imprecision in the specification of a parametric analysis and enhances at the same time the flexibility of the analysis.

In Bayesian inference an imprecise input is typically represented by a class of probability measures on the unknown quantities. In the previous literature about parametric robustness most of the efforts have been concentrated on classes of prior distributions for a finite collection of unknown parameters, while the distribution of the data conditionally on those parameters is considered to be known exactly. Robustness with respect to misspecification has received less attention due to the mathematical difficulties. In the nonparametric context the two aspects are no longer separated but they are in fact identified with the presence of a unique unknown infinite-dimensional parameter.

We examine here nonparametric analysis within the context of Exchangeable Tree processes that fall in the general class of exchangeable processes and represent a subclass of Tail-free processes. Exchangeable Trees constitute indeed a general class of processes and contain as particular cases Dirichlet processes and Polya Trees, two of the most used nonparametric priors.

We propose a predictive interpretation for an imprecise prior input that leads us to formulate a general solution for the global robustness investigation. We are then able to quantify the range of linear functionals of the conditional predictive distribution after some data have been collected.

The larger framework implied by enlarging the scope to an infinite-dimensional parameter leads us to expect less robustness. Some annoying phenomena, like dilation, are experienced, deviating from the usual pattern in parametric robustness.

We are able to compare how this is affected by the prior inputs and quantify how robustness can be improved by restricting attention to particular subclasses.

Finally, a different problem is approached: simulation from mixture distributions whose components are supported on spaces of different dimensions. Here a novel approach is considered by reducing the problem to that of simulating from a single target distribution that is absolutely continuous with respect to the Lebesgue measure on the largest support of the components. This approach is suggested from an alternative representation of the simulation goal in the simplified situations when the mixture consists only of a one-dimensional component and a degenerate component. Tools for suitable generalization to arbitrarily nested components and to an arbitrary number of them are provided. Hence two alternative methods are derived and one of them is successfully employed in analyzing simulated data as well as a real data set.

The new approach is designed in order to avoid the numerical integration needed for evaluating the relative weight of each component and represents in the case of nested components an alternative to the currently available Monte Carlo Markov Chain methods such as the reversible jump algorithm (Green, 1995) and the composite-space approach (Carlin and Chib, 1995). The distinguishing feature of the proposed method is the absence of proposals for jumping between components of different dimension or of the specification of pseudopriors. This allows for a more automatic implementation. Furthermore it is argued that in the actual implementation of a Markov chain that simulates from the absolutely continuous target distribution one can automatically build up a chain that allows for moving from one component to any other possible component which possibly improves the speed of convergence. Finally, in order to assess the mixing behaviour, standard convergence diagnostics for absolutely continuous stationary distributions can be used.