Posterior contraction
for conditionally Gaussian priors

René de Jonge

In this thesis we consider two examples of conditionally Gaussian stochastic processes for the construction of prior distributions on statistical models indexed by an infinite dimensional parameter in a space of functions, for example a probability density or a regression function. We allow these functions to be functions of several variables, so the stochastic processes in question might in fact be stochastic fields. We moreover assume that these functions are smooth in the sense of Hölder. The two examples that we construct are defined by choosing the paths of the process to be either tensor-product spline functions or location-scale kernel mixtures.

The use of log spline models and kernel mixtures to construct priors on probability densities is well-established in Bayesian nonparametrics. The use of Gaussian priors provides a unified approach to obtain posterior contraction rates in various statistical settings simultaneously.

Our goal is to obtain an adaptive procedure with consistent and rate-optimal posterior contraction. If the function to be estimated is a function of $d$ variables with smoothness level $\alpha$, then the typical optimal rate of posterior contraction is of the order $n^{-\frac{\alpha}{d+2}}$, with $n$ the number of observations. We show that it is possible to construct Gaussian priors from either the spline functions or the kernel mixtures which actually achieve posterior contraction at a near optimal rate. These priors will however depend on $\alpha$, the smoothness of the function to be estimated. We show that in both cases it is possible to define a hierarchical procedure based on these Gaussian priors which also achieves a near optimal rate of posterior contraction, but which itself does not depend on the level of smoothness of the function to be estimated, so that this procedure adapts to the smoothness.

In the setting of fixed design regression with Gaussian errors, the variance of the errors is a finite dimensional nuisance parameter which we can equip with a prior as well. The posterior contraction results also imply the concentration of posterior mass around this finite dimensional parameter at the nonparametric rate. We however know that posterior contraction in the finite parameter case is typically faster: the optimal rate in is $n^{-1/2}$. We show via a semi-parametric Bernstein-von Mises result that it is possible to achieve posterior contraction around the finite dimensional parameter at rate $n^{-1/2}$ if we equip the infinite dimensional parameter, the regression function $f$, as before with a Gaussian prior distribution.