Convergence

McMc methods require the use of diagnostics to assess whether the iterative simulations have reached the equilibrium distribution of the markov chain. Sampled chains require to be run for an intial burn-in period until they can be assumed to provide approximately correct samples from the posterior distribution of interest. This burn -in period can vary considerably between different problems. In addition, it is important to ensure that the chain manages to explore the parameter space properly so that the sampler does not ‘stick’ in local maxima of the surface of the distribution. Hence, it is crucial to ensure that a burn-in period is adequate for the problem considered. Judging convergence has been the subject of much debate and can still be regarded as art rather than science: a qualitative judgement has to be made at some stage as to whether the burn-in period is long enough..

There are a wide variety of methods now

available to assess convergence of chains within McMC. cite: robcasella and cite: chenib provide recent reviews. The available methods are largely based on checking the distributional properties of samples from the chains.

Single chain methods

First, global methods for assessing convergence have been proposed which involve monitoring functions of the posterior probability.at each iteration. These methods look or stabilisiation of the probability value. This value forms a time series, and special cusum methods have been proposed (cite: yumyk).

Second, graphical methods have been proposed which allow the comparison of the whole distribution of successive samples. Quantile-quantile plots of successive lengths of single variable output from the sampler can be used for this purpose.

Multi-chain methods

Single chain methods can, of course, be applied to each of a multiple of chains. In addition, there are methods that can only be used for multiple chains. The Gelman-Rubin statistic was proposed as a method for assessing the convergence of multiple chains via the comparison of summary measures across chains (cite: gelrubin, cite: brookgel, cite: robcasella, Ch. 8). There is some debate about whether it is useful to run one long chain as opposed to multiple chains with different start points. The advantage of multiple chains is that they provide evidence for the robustness of convergence across different subspaces. However, as long as a single chain samples the parameter space adequately, then these have benefits. The reader is referred to cite: robcasella, chapter 8 for a thorough discussion of diagnostics and their use.