RSS Ordinary meeting, 20 October 1993

Contribution to discussion of paper by Grenander and Miller

P. J. Green, University of Bristol

I shall comment here only on the jump-diffusion dynamics, providing the computational engine for the fascinating vision applications in the paper.

Gibbs-sampler-like dynamics like those of Corollary 1 will rarely be useful; Markov chain simulation is unnecessary if jumping into a subspace in equilibrium is possible. Much more generally, jumps must be controlled by Metropolis-Hastings acceptance-rejection decisions, and here I will describe a framework for doing this, discussing only the reversible case. Sufficient conditions are given in Theorem 1, but the authors' specific instances of jump measures provide little guidance for the reader with a new problem.

Working in discrete time, consider a family of possible "moves", indexed by m, counting, for example, a birth from k to $k+1$ objects in the mitochondria example the same move as a death from $k+1$ to k. When the current state is $x \in \mathcal{C}$, the probability of proposing a move of type m and landing in dy is $r_m(x, dy)$.

Some dimension-matching is necessary. Writing ν_m for the measure on $\mathcal{C} \times \mathcal{C}$ given by

$$\nu_m(A \times B) = \int_{x \in A} \int_{y \in B} \mu(dx) r_m(x, dy),$$

suppose that for each m, $\nu_m(A \times B)$ and $\nu_m(B \times A)$ have finite positive densities with respect to the same *symmetric* measure ξ_m on $\mathcal{C} \times \mathcal{C}$.

Then

$$\alpha_m(x, y) = \min \left\{ 1, \frac{\mu(dy) r_m(y, dx)}{\mu(dx) r_m(x, dy)} \right\}$$

is well-defined (the ratio is that of the two afore-mentioned densities). This is similar to the usual Hastings ratio, though here the r_m are improper distributions, and on different subspaces. The proposed move is accepted with probability $\alpha_m(x, y)$; otherwise x is retained. The resulting jump measure is

$$q(x, dy) = \sum_m \alpha_m(x, y) r_m(x, dy)$$

Moves within subspaces can be expressed in the same notation (although I do not find the diffusion idea here very helpful; in practice, time will be discretised into positive increments, and I would prefer to work in discrete time intentionally, using a process exactly preserving the required equilibrium).

It is clear that the resulting chain satisfies the detailed balance condition

$$\int_A \mu(dx) \int_B P(x \to dy) = \int_B \mu(dy) \int_A P(y \to dx).$$

This framework covers fixed-parameter-space MCMC, as used in general Bayesian inference and pixel-based imaging, the birth-death simulations used in point processes, and the examples in the paper.

For a simple concrete example, suppose the target distribution has just two components, with probabilities p_1, p_2, a univariate density $\pi_1(x)$ on $[0,1]$ and a bivariate density $\pi_2(x_1, x_2)$ on $[0,1] \times [0,1]$. We might consider proposals summarised by:

Move type	$x \to$	$(x_1, x_2) \to$
1	(x, u)	x_1
2	(u, x)	x_2
3	(u_1, u_2)	u
4	$(x-u, x+u)$	$\frac{1}{2}(x_1+x_2)$

where the u's are independently uniformly distributed on the appropriate intervals. If the probabilities of each move type are bounded away from zero, the dimension-matching requirement is met, although the proposal densities have dimensions not determined solely by those of the takeoff and landing subspaces. The $\alpha_m(x, y)$ are easily derived.

I will conclude with some miscellaneous comments. These methods require μ_k to be properly normalised within each subspace; the $Z(k)$ (or at least their ratios) must be known. When this is impossible, e.g. in Markov random fields, experiments we have done using stochastic approximation ideas may be generalisable to allow learning of $Z(k)$ during the course of the simulation.

The context of our experiments was *simulated tempering* (Marinari and Parisi, 1992), and as there one could contemplate the invention of additional subspaces and associated models, simply to facilitate mixing.

Finally, applications such as those discussed here highlight the great need for new diagnostic and design tools for MCMC, to establish and monitor convergence.

Reference

Marinari, E. and Parisi, G. (1992) Simulated tempering: a new Monte Carlo scheme. *Europhys. Lett.*, **19**, pp. 451–458.