Discussion of the Read Paper by Girolami and Calderhead “Riemann manifold Langevin and Hamiltonian Monte Carlo methods” read to the Society on October 13th, 2010

Magali Beffy and Christian P. Robert

CREST, Paris, and Université Paris–Dauphine, CEREMADE

The paper gives a very clear geometric motivation for the use of an Hamiltonian representation. As such, it suggests for an immediate generalization by extending the Hamiltonian dynamic to a more general dynamic on the level sets of different Hamiltonians

$$\mathscr{H}(\theta, \mathbf{p}_k) = -\mathcal{L}(\theta) + \frac{1}{2}\log\{(2\pi)^D|\mathbf{G}_k(\theta)|\} + \frac{1}{2}\mathbf{p}_k^{\mathrm{T}}\mathbf{G}_k(\theta)^{-1}\mathbf{p}_k$$

where the $\mathbf{p}_j$s are auxiliary vectors of the same dimension D as θ and the $\mathbf{G}_j(\theta)$s are symmetric matrices. This function is then associated with the pde's

$$\begin{aligned}\frac{\mathrm{d}\theta_i}{\mathrm{dt}} &= \frac{\partial}{\partial p_{ij}}\mathscr{H}(\theta, \mathbf{p}_j) = \left\{G_j(\theta)^{-1}\mathbf{p}_j\right\}_i \\ \frac{\mathrm{d}p_{ij}}{\mathrm{dt}} &= -\frac{\partial}{\partial \theta_i}\mathscr{H}_j(\theta, \mathbf{p}_j)\end{aligned}$$

The target distribution is preserved at all times t. This generalization would allow for using a range of information matrices $\mathbf{G}_j(\theta)$s in parallel. The corresponding RHMC implementation is to pick one of the indices j at random and to follow the same moves as in the paper, given the separation between the different energies.