Hoffman Gelman 2014
The No-U-Turn Sampler

Daniele Zago

May 13, 2021

1 Summary

Standard HMC

As MCMC and Gibbs sampling are particularly inefficient, *Hamiltonian Monte Carlo* (HMC) (Algorithm 1) is able to improve the sampling performance by suppressing their random walk behaviour via the introduction of Hamiltonian dynamics. For a distribution in $\mathbb{R}^D$, HMC has a cost of $\mathcal{O}(D^{5/4})$ in comparison to $\mathcal{O}(D^2)$ of MCMC. The *No-U-Turn Sampler* (NUTS) is an improvement of HMC which removes the need of specifying the number of "leapfrog" steps L for which to run the Hamiltonian system.

No-U-Turn Sampler

The leapfrog steps are run until the dot product between $\tilde{r}$ (current momentum) and $\tilde{\vartheta} - \vartheta$ (diff. btw initial position and current position) is negative, i.e. we are not moving away. The steps are taken in a binary tree fashion and then the update is carefully sampled from the generated set of points, in order to preserve time-reversibility of the algorithm and detailed balance, therefore guaranteeing that the MCMC update is correct.

Algorithm 1 Standard Hamiltonian Monte Carlo

for $m = 1$ to M **do**
 Sample $r^0 \sim \mathcal{N}(0, I)$ ▷ Sample the momentum variable
 Set $\vartheta^m \leftarrow \vartheta^{m-1}$, $\tilde{\vartheta} \leftarrow \vartheta^{m-1}$, $\tilde{r} \leftarrow r^0$
 for $i = 1$ to L **do** ▷ Perform L leapfrogs
  $\tilde{\vartheta}, \tilde{r} \leftarrow \text{Leapfrog}(\tilde{\vartheta}, \tilde{r}, \varepsilon)$
  Set $\vartheta^m \leftarrow \tilde{\vartheta}, r^m \leftarrow -\tilde{r}$ with prob. $\alpha = \min\left\{1, \frac{\exp\left\{\mathcal{L}(\tilde{\vartheta}) - \frac{1}{2}\tilde{r}\cdot\tilde{r}\right\}}{\exp\left\{\mathcal{L}(\vartheta^{m-1}) - \frac{1}{2}r^0\cdot r^0\right\}}\right\}$
 end for
end for
function $\text{Leapfrog}(\vartheta, r, \varepsilon)$
 $\tilde{r} \leftarrow r + \varepsilon/2 \cdot \nabla_\vartheta \mathcal{L}(\vartheta)$
 $\tilde{\vartheta} \leftarrow \vartheta + \varepsilon\tilde{r}$
 $\tilde{r} \leftarrow r + \varepsilon/2 \cdot \nabla_\vartheta \mathcal{L}(\tilde{\vartheta})$
end function

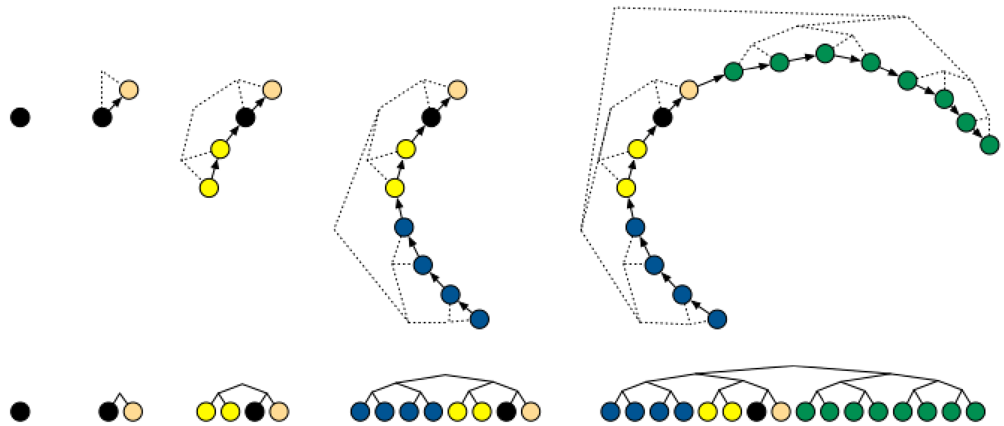

Figure 1: The NUTS sampler uses a binary tree expansion until the particle turns back on itself, and then samples from the obtained locations.