Hamiltonian Monte Carlo for Binary Distributions with Augmentation

Yuze Zhou

Sep.2020

Augmentation for Binary Random Vector

- Let $s \in \{-1, +1\}^d$ be the d-dimensional binary random vectors with probability mass function:

$$p(s) \propto f(s)$$

- We then augment s with continuous variable y such that conditional distribution $p(y|s)$ only has density at the orthant where:

$$s_i = sign(y_i)$$

- The potential energy function $U(y)$ is therefore:

$$U(y) = -log(p(y|s)) - log(f(s))$$

- The Hamiltonian is therefore:

$$H(y, q) = U(y) + \frac{1}{2}\|q\|_2^2$$

Hamiltonian Dynamics for Augmented Distribution

- When staying in the same orthant, the dynamics would move according to the Hamiltonian equations as:

$$\frac{dy}{dt} = \frac{\partial U}{\partial q}$$
$$\frac{dq}{dt} = -\frac{\partial U}{\partial y}$$

- The zero divergence of the Hamiltonian ensures that the Hamiltonian stays constant while in each orthant, but it would get interrupted when some coordinate hits zero where the potential energy function $U(y)$ becomes discontinuous.

Hamiltonian Dynamics for Augmented Distribution

- Conservation of the Hamiltonian imposes a jump in q_j when discontinuity happens to y_j. Assume that the jump happens to y_j at time t_j from $y_j < 0$ to $y_j > 0$.
- Let $q_j(t_j-)$ and $q_j(t_j+)$ be the value of the momentum q_j just before and after t_j
- To enforce conservation of energy, we equate the Hamiltonian at both sides, and will yield:

$$\frac{q_j^2(t_j+)}{2} = \Delta_j + \frac{q_j^2(t_j-)}{2}$$
$$\Delta_j = U(y_j = 0, s_j = +1) - U(y_j = 0, s_j = -1)$$

- If the right hand side yields to a positive value for $q_j(t_j+)$, we change the sign of both y_j and s_j right after t_j, but $q_j(t_j+)$ will have the same sign as $q_j(t_j-)$.
- If the right hand side yields to a negative value for $q_j(t_j+)$, we keep the sign for both y_j and s_j right after t_j, but $q_j(t_j+)$ will change sign compared to $q_j(t_j-)$.
- For T large enough, we sample $sign(y_i(T)) = s_i$ as a candidate for s

Gaussian Augmentation

- The binary distribution $f(s)$ is augmented with truncated Gaussians as:

$$p(y|s) = \begin{cases} (\frac{2}{\pi})^{\frac{d}{2}} exp(-\frac{\|y\|_2^2}{2}) & when \quad sign(y_i) = s_i \\ 0 & otherwise \end{cases}$$

- The truncated Gaussian will yield to that:

$$\frac{d^2 y}{dt^2} = -y$$
$$\frac{d^2 q}{dt^2} = -q$$

which has a solution:

$$y_i(t) = y_i(0)cos(t) + q_i(0)sin(t)$$
$$q_i(t) = -y_i(0)sin(t) + q_i(0)cos(t)$$

Gaussian Augmentation

- The differential equations above also indicates that after y_j reaches the boundary 0 at time t_j, and the new $q_j(t_j+)$ is obtained, the trajectories of the dynamics after t_j (until the boundary is hit again) would be:

$$y_j(t) = q_j(t_j+)sin(t - t_j)$$
$$q_j(t) = q_j(t_j+)cos(t - t_j)$$

- The solutions above indicates that the hitting times would have a periodic behaviour, let $0 < t_{j_1} < t_{j_2} < \cdots < t_{j_d} < \pi$ be all the d initial hitting times in all d coordinates, all the other incoming hitting times would therefore be:

$$t_i + n\pi \quad n \in \mathbb{N}$$

Exponential Augmentation

- Here we augment $f(s)$ with:

$$p(y|s) = \begin{cases} exp(-\sum_{i=1}^{d} |y_i|) \quad when \quad sign(y_i) = s_i \\ 0 \quad otherwise \end{cases}$$

- The Hamiltonian equations (when staying in the same orthant) now satisfies that:

$$\frac{d^2 y}{dt^2} = -s$$

which has solutions:

$$y_j(t) = y_i(0) + q_i(0)t - \frac{s_j t^2}{2}$$

Exponential Augmentation

- When y_j hits the boundary 0 at t_j and we update $q_j(t_j+)$ as aforementioned, the dynamics continues in the jth coordinate after t_j as:
$$y_j(t) = q(t_j+)(t - t_j) - \frac{s_j}{2}(t - t_j)^2$$
- The next hitting time in the j the coordinate would therefore be:
$$t_j + 2|q_j(t_j+)|$$

References I