Семинар 9. Гибридный метод Монте-Карло (HMC)

Курс: Байесовские методы в машинном обучении, 2017

1. Докажите эквивалентность системы
$$\begin{cases} \frac{\mathrm{d}\theta}{\mathrm{d}t} & = \mathbf{v} \\ m\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} & = -\nabla_\theta U(\theta) \end{cases}$$
системе уравнений Гамильтона
$$\begin{cases} \frac{\mathrm{d}\theta}{\mathrm{d}t} = \nabla_r H(\theta, r) \\ \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = -\nabla_\theta H(\theta, r), \end{cases}$$
где $H(\theta, r) = U(\theta) + \frac{1}{2m} r^T r$ и $r = mv$.

2. Для решений уравнений Гамильтона $\theta(t), r(t)$ проверьте закон сохранения энергии
$$\frac{\mathrm{d}H(\theta(t), \mathbf{r}(\mathbf{t}))}{\mathrm{d}t} = 0$$
и свойство задаваемого решениями векторного поля $V(t) = (\frac{\mathrm{d}\theta}{\mathrm{d}t}, \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t})$
$$\operatorname{div} V(t) = 0.$$

3. Для вероятностного распределения $p(\theta) = \frac{1}{Z}\exp(-\cos\theta) I[\theta \in (-4\pi, 4\pi)]$ выписать соответствующий Гамильтониан $H(\theta, r)$, нарисовать его линии уровня. Как они связаны с траекториями решения уравнений Гамильтона?

4. Рассмотрим вероятностной распределение вида $p(\theta) = \frac{1}{Z}\exp(-U(\theta))$ и соответствующие ему уравнения Гамильтона. Определим для некотороых моментов времени t_0 и t_1 на фазовом пространстве преобразование T по правилу
$$T(\theta, \mathbf{r}) = (\theta(t_1), -\mathbf{r}(t_1)),$$
где $\theta(t)$ и $r(t)$ - решения уравнений Гамильтона с начальными условиями $\theta(t_0) = \theta$, $r(t_0) = r$. Докажите, что
$$T(T(\theta, \mathbf{r})) = (\theta, \mathbf{r})$$
и проиллюстрируйте это на примере из прошлой задачи. Покажите, что при использовании в качестве предложного распределения в алгоритме Метрополиса-Хастингса распределения
$$q(\theta', \mathbf{r}' | \theta, \mathbf{r}) = \delta\left((\theta', \mathbf{r}') - T(\theta, \mathbf{r})\right)$$
вероятность принятия всегда будет равна единице.

5. Выпишите итерационную схему алгоритма Hamiltonial Monte-Carlo для кинетической энергии общего вида $K(\mathbf{r}) = \frac{1}{2} r^T M^{-1} r$. Покажите, что для $M = I$ при использовании лишь одного шага итерации алоритм совпадает с динамикой Ланжевена.