

## Начальные распределения для задачи двух тел

### *MCMC-sampling*

Предположим мы генерируем последовательность случайных величин,  $\{X_0, X_1, X_2, \dots\}$ , такую что в каждый момент  $t \geq 0$  следующее состояние  $X_{t+1}$  выбирается исходя из распределения  $P(X_{t+1}|X_t)$ , которое зависит от текущего состояния  $X_t$ , но не от предыдущего набора состояний  $\{X_0, X_1, X_2, \dots, X_{t-1}\}$ . То есть, состояние  $X_{t+1}$  определяется исключительно предыдущим  $X_t$ . Такая последовательность состояний называется *цепью Маркова*.

Рассмотрим алгоритм Метрополиса-Гастингса, позволяющий получать последовательность точек – элементов Марковской цепи – распределенную согласно заданной плотности вероятности  $\pi(\cdot)$ .

---

#### **Algorithm 1** Metropolis-Hastings algorithm [1]

---

```

1: Initialize  $x^{(0)} \sim q(x)$ 
2: for iteration  $i = 1, 2, \dots$  do
3:   Propose:  $x^{cand} \sim q(x^{(i)}|x^{(i-1)})$ 
4:   Acceptance probability:
5:      $\alpha(x^{cand}|x^{(i-1)}) = \min \left\{ 1, \frac{q(x^{(i-1)}|x^{cand})\pi(x^{cand})}{q(x^{cand}|x^{(i-1)})\pi(x^{(i-1)})} \right\}$ 
6:    $u \sim \text{Uniform}(0, 1)$ 
7:   if  $u < \alpha$  then
8:     Accept the proposal:  $x^{(i)} \leftarrow x^{cand}$ 
9:   else
10:    Reject the proposal:  $x^{(i)} \leftarrow x^{(i-1)}$ 
11:   end if
12: end for
```

---

Первым шагом алгоритма является выбор случайной точки (эта величина выбирается определенным образом на основе распределения; я же выбирал ее совершенно случайным образом, но так, чтобы она не оказалась в какой-то физически маловероятной области). Следующий за ним главный цикл алгоритма состоит из трех частей: (1) Получать следующую точку ("кандидата")  $x^{cand}$  исходя из вспомогательного распределения  $q(x^{(i)}|x^{(i-1)})$ ; (2) Рассчитать вероятность перехода в новую точку  $\alpha(x^{cand}|x^{(i-1)})$ , основываясь на распределении  $q$  и функции распределения  $\pi$ ; (3) Принять новую точку с вероятностью  $\alpha$ .

Обратим внимание на то, что точка, полученная исходя из вспомогательного распределения  $q(\cdot)$ , принимается не всегда, а лишь с вероятностью  $\alpha(\cdot)$ . Рассматривают вспомогательные распределения двух классов – симметричные и асимметричные. Симметричным называется распределение, удовлетворяющее следующему соотношению

$$q(x^{(i)}|x^{(i-1)}) = q(x^{(i-1)}|x^{(i)})$$

К часто используемым симметричным распределениям относятся гауссово и равномерное распределения. В качестве примера рассмотрим вспомогательное распределение Гаусса:

$$x^{cand} = x^{(i-1)} + \text{Normal}(0, \sigma)$$

Понятно, что  $\text{Normal}(x^{cand} - x^{(i-1)}; 0, \sigma) = \text{Normal}(x^{(i-1)} - x^{cand}; 0, \sigma)$ , то есть Гауссово распределение в действительности задает симметричное вспомогательное распределение.

Среднеквадратичное отклонение  $\sigma$  является параметром модели. Значение этого параметра будет определять динамику Марковской цепи в рассматриваемом пространстве.

В случае симметричных вспомогательных распределений выражение для вероятности выбора новой точки  $\alpha(\cdot)$  существенно упрощается:

$$\alpha(x^{cand}|x^{(i-1)}) = \min \left\{ 1, \frac{\pi(x^{cand})}{\pi(x^{(i-1)})} \right\}$$

Заметим, что если плотность вероятности (точнее говоря, величина, пропорциональная плотности вероятности) в новой точке  $\pi(x^{cand})$  больше, чем плотность вероятности в текущей  $\pi(x^{(i-1)})$ , то их отношение будет больше 1, а значит вероятность перехода в новую точку будет равна 1:  $\alpha(x^{cand}|x^{(i-1)}) = 1$ . Другими словами, если новая точка выбрана таким образом, что плотность вероятности в ней больше, чем в текущей, то в нее осуществляется переход. Устройство алгоритма таково, что Марковская цепь "склонна" посещать те точки пространства, в которых моделируемая плотность вероятности выше. Однако, если новая точка была выбрана таким образом, что плотность вероятности в ней меньше, чем в текущей, то тогда вероятность перейти в нее будет определяться отношением плотностей вероятности:

$$\alpha(x^{cand}|x^{(i-1)}) = \frac{\pi(x^{cand})}{\pi(x^{(i-1)})}$$

То есть, если вероятность в новой точке будет мала по сравнению с текущей, то и переход в нее будет маловероятен.

Вид вероятности перехода в новую точку из текущей определяется *условием детального баланса* [2]. Последнее гарантирует, что полученная Марковская цепь в действительности будет удовлетворять заданной плотности вероятности.

### Точные формулы для двухатомной системы

Рассмотрим вектор, соединяющий центры атомов. Обозначим  $\mathbf{r}$  его координаты в лабораторной системе координат,  $\mathbf{R}$  – в молекулярной системе координат. Производные  $\dot{\mathbf{r}}$  и  $\dot{\mathbf{R}}$  связаны при помощи матрицы эйлеровых углов  $\mathbb{S}$  и угловой скорости  $\boldsymbol{\Omega}$ :

$$\dot{\mathbf{r}} = \mathbb{S}^{-1} \left( \dot{\mathbf{R}} + [\boldsymbol{\Omega} \times \mathbf{R}] \right). \quad (1)$$

Пусть атомы в молекулярной системе координат расположены на оси  $Z$ , в таком случае правая часть выражения (1) превращается в

$$\begin{aligned} \dot{\mathbf{r}} &= \mathbb{S}^{-1} \left\{ \begin{bmatrix} 0 \\ 0 \\ \dot{R} \end{bmatrix} + \begin{bmatrix} \Omega_y R \\ -\Omega_x R \\ 0 \end{bmatrix} \right\} \\ \mathbb{S} \dot{\mathbf{r}} &= \begin{bmatrix} \Omega_y R \\ -\Omega_x R \\ \dot{R} \end{bmatrix}. \end{aligned} \quad (2)$$

Лагранжиан в молекулярной системе координат имеет следующий вид:

$$\mathcal{L} = \frac{1}{2} \mu \dot{R}^2 + \frac{1}{2} \boldsymbol{\Omega}^\top \begin{bmatrix} \mu R^2 & 0 & 0 \\ 0 & \mu R^2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \boldsymbol{\Omega}$$

Используя теорему Донкина, находим связь гамильтоновых переменных  $\mathbf{J}$  и  $\mathbf{p} = [p_R]$  с лагранжевыми переменными  $\boldsymbol{\Omega}$  и  $\mathbf{q} = [R]$ :

$$\begin{aligned} \mathbf{J} = \frac{\partial \mathcal{L}}{\partial \boldsymbol{\Omega}} = \boldsymbol{\Omega} & \quad J_x = \mu R^2 \Omega_x \\ \mathbf{p} = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} = \dot{\mathbf{q}} & \quad J_y = \mu R^2 \Omega_y \\ & \quad p_R = \mu \dot{R} \end{aligned} \quad \Longrightarrow \quad (3)$$

Выкладка в приложении А показывает, что каждая компонента  $\dot{\mathbf{r}}$  имеет нормальное распределение  $\dot{\mathbf{r}} \sim \mathcal{N}\left(0, \frac{kT}{\mu}\right)$ . Таким образом,

**Литература**

1. Yildirim I. Bayesian Inference: Metropolis-Hastings Sampling. MIT Online Library
2. Gilks, W.R., Richardson, S., & Spiegelhalter, D.J. (1996). *Markov Chain Monte Carlo in Practice*. London: Chapman and Hall.

---

# Appendices

## Приложение А. Распределения в лабораторной системе координат

Воспользуемся следующими двумя выводами из теории вероятностей:

1. Пусть случайная величина  $\xi$  распределена с плотностью  $f_\xi(x)$ . Тогда случайная величина  $\eta = a\xi + b$  распределена с плотностью

$$f_\eta(x) = \frac{1}{|a|} f_\xi\left(\frac{x-b}{a}\right)$$

2. Если две независимые случайные величины  $X$  и  $Y$  распределены с плотностями  $X \sim f_1(x)$  и  $Y \sim f_2(x)$  соответственно, то случайная величина  $Z = X + Y$  распределена с плотностью

$$g(z) = \int_{-\infty}^{+\infty} f_1(x) f_2(z-x) dx$$

Т.к. вектор  $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$  равен разнице радиус-векторов двух атомов  $\mathbf{r}_1$  и  $\mathbf{r}_2$  в лабораторной системе координат соответственно, то  $\dot{\mathbf{r}} = \dot{\mathbf{r}}_1 - \dot{\mathbf{r}}_2$ . Используя п.1 и п.2 получим распределение для компонент  $\mathbf{r}$ :

$$\begin{cases} \dot{\mathbf{r}}_{1x} \sim f_1(x) = \sqrt{\frac{m_1}{2\pi kT}} \exp\left(-\frac{m_1 x^2}{2kT}\right) \\ -\dot{\mathbf{r}}_{2x} \sim f_2(x) = \sqrt{\frac{m_2}{2\pi kT}} \exp\left(-\frac{m_2 x^2}{2kT}\right) \end{cases}$$
$$\dot{\mathbf{r}}_x \sim \int_{-\infty}^{+\infty} f_1(x) f_2(z-x) dx = \frac{\sqrt{m_1 m_2}}{2\pi kT} \int_{-\infty}^{+\infty} \exp\left(-\frac{m_1 x^2}{2kT}\right) \exp\left(-\frac{m_2 (z-x)^2}{2kT}\right) dx \quad (4)$$

Отдельно рассмотрим получившийся интеграл:

$$\begin{aligned} \int_{-\infty}^{+\infty} \exp\left(-\frac{m_1 x^2}{2kT} - \frac{m_2 (z-x)^2}{2kT}\right) dx &= \int_{-\infty}^{+\infty} \exp\left(-\frac{(m_1 + m_2) x^2 - 2m_2 z x + m_2 z^2}{2kT}\right) dx = \\ &= \int_{-\infty}^{+\infty} \exp\left(-\frac{\left(\sqrt{m_1 + m_2} x - \frac{m_2}{\sqrt{m_1 + m_2}} z\right)^2}{2kT}\right) \exp\left(-\frac{m_2 z^2 - \frac{m_2^2}{m_1 + m_2} z^2}{2kT}\right) dx = \\ &= \left[ y = \frac{\sqrt{m_1 + m_2} x - \frac{m_2}{\sqrt{m_1 + m_2}} z}{\sqrt{2kT}} \right] = \sqrt{\frac{2kT}{m_1 + m_2}} \exp\left(-\frac{m_1 m_2}{2(m_1 + m_2) kT} z^2\right) \int_{-\infty}^{+\infty} \exp(-y^2) dy = \\ &= \sqrt{\frac{2\pi kT}{m_1 + m_2}} \exp\left(-\frac{m_1 m_2}{2(m_1 + m_2) kT} z^2\right) \end{aligned} \quad (5)$$

Подставляя значение интеграла (5) в выражение для плотности распределения  $\dot{\mathbf{r}}_x$  (4), получаем

$$\dot{\mathbf{r}}_x \sim \frac{1}{\sqrt{2\pi kT}} \sqrt{\frac{m_1 m_2}{m_1 + m_2}} \exp\left(-\frac{m_1 m_2}{2(m_1 + m_2) kT} z^2\right) = \sqrt{\frac{\mu}{2\pi kT}} \exp\left(-\frac{\mu z^2}{2kT}\right),$$

где через  $\mu$  была обозначена приведенная масса двухатомной системы  $\mu = \frac{m_1 m_2}{m_1 + m_2}$ .