1 Начальные распределения для задачи двух тел

1.1 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 $P\left(X_{t+1}|X_t\right)$, которое зависит от текущего состояния X_t, но не от предыдущего набора состояний $\{X_0, X_1, X_2 ... X_{t-1}\}$. 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]

Initialize $x^{(0)} \sim q(x)$
for iteration $i = 1, 2, \dots$ **do**
  Propose: $x^{cand} \sim q\left(x^{(i)}|x^{(i-1)}\right)$
  Acceptance probability:
    $\alpha\left(x^{cand}|x^{(i-1)}\right) = \min\left\{1, \frac{q\left(x^{(i-1)}|x^{cand}\right)\pi\left(x^{(cand)}\right)}{q\left(x^{cand}|x^{(i-1)}\right)\pi\left(x^{(i-1)}\right)}\right\}$
  $u \sim$ Uniform(u; 0, 1)
  if $u < \alpha$ **then**
    Accept the proposal: $x^{(i)} \leftarrow x^{cand}$
  else
    Reject the proposal: $X^{(i)} \leftarrow x^{(i-1)}$
  end if
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 $q\left(x^{(i)}|x^{(i-1)}\right)$; (2) Рассчитать вероятность перехода в новую точку $\alpha\left(x^{cand}|x^{(i-1)}\right)$, основываясь на распределении q и функции распределения π; (3) Принять новую точку с вероятностью α.

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

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

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

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

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

Среднеквадратичное отклонение σ 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:

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

Заметим, что если плотность вероятности (точнее говоря, величина, пропорциональная плотности вероятности) в новой точке $\pi\left(x^{cand}\right)$ больше, чем плотность вероятности в текущей $\pi\left(x^{(i-1)}\right)$, то их отношение будет больше 1, а значит вероятность перехода в новую точку будет равна 1: $\alpha\left(x^{cand}|x^{(i-1)}\right) = 1$. 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:

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

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.

2 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.