

Домашнее задание 8. MCMC

Курс: Байесовские методы в анализе данных, 2019

Задача 1. Запишите алгоритм генерации $x \sim p(x) = C(x|a, b) = \frac{1}{\pi b[1+(\frac{x-a}{b})^2]}$ (распределение Коши) с помощью метода обратной функции.

Задача 2. Метод Stochastic gradient Langevin dynamics (SGLD) — один из простых примеров применения Markov Chain Monte Carlo к нейронным сетям. SGLD обновляет веса нейронной сети по следующей формуле:

$$w_{n+1} = w_n + \frac{\epsilon}{2} \left(\nabla \log p(w_n) + \frac{N}{n} \sum_{i=1}^n \nabla \log p(y_{j_i} | x_{j_i}, w_n) \right) + \eta_n, \quad \eta_n \sim \mathcal{N}(\eta | 0, \epsilon I), \quad (1)$$

где

- $j_i \sim \text{Unif}(1, \dots, N)$ — индексы объектов мини-батча, n — размер мини-батча, N — размер обучающей выборки;
- $p(w)$ — априорное распределение на веса, $p(y|x, w)$ — правдоподобие данных;
- ϵ — длина шага.

. Процедура очень похожа на стохастический градиентный спуск (SGD), но имеет два отличия: зашумление вектора стохастического градиента функционала качества (η_n) и пропуск некоторых семплов согласно тесту Метрополиса-Хастингса. Благодаря зашумлению, метод не просто сходится к локальному максимуму апостериорного распределения на веса (как это делает SGD), а генерирует веса в окрестности максимума, а также может «переходить» из окрестности одного максимума в окрестность другого максимума.

Задания:

- Запишите предположное распределение $r(\cdot | w_n)$, соответствующее формуле перехода 1.
- Является ли предположное распределение симметричным ($r(a|b) = r(b|a)$)? Поясните ответ.
- Запишите детальную схему Метрополиса-Хастингса для метода SGLD (из распределений в ней могут присутствовать только $p(y|x, w)$, $p(w)$, $r(\cdot | w_n)$).

Задача 3. Предположим, что выборка $X = (x_1, \dots, x_N)$ сгенерирована из экспоненциального распределения:

$$p(X|a, b) = \prod_{i=1}^N p(x_i|a, b), \quad p(x|a, b) = \text{Exp}(x|ab) = ab \exp(-abx), \quad x > 0.$$

Введем экспоненциальное априорное распределение на параметры модели:

$$p(a, b) = \exp(-a - b), \quad a > 0, \quad b > 0.$$

Запишите схему Гиббса для семплирования параметров a, b из апостериорного распределения $p(a, b|X)$. Найдите условные распределения, фигурирующие в схеме (придется вспомнить условное сопряжение распределений :)