Markov Chain Monte Carlo (MCMC)

Михаил Ураков

8 ноября 2025 г.

План

В чём проблема?

Сложные интегралы

Во многих областях мы сталкиваемся с необходимостью вычисления сложных многомерных интегралов:

$$\mathbb{E}[f(\theta)] = \int f(\theta)p(\theta)d\theta$$

где $p(\theta)$ – апостериорное распределение.

Почему это сложно?

- Много измерений (curse of dimensionality)
- Нет аналитического решения
- Невозможно напрямую сэмплировать из $p(\theta)$

Monte Carlo: Интуиция

Основная идея

Заменяем интегрирование на усреднение по выборке:

$$\mathbb{E}[f(\theta)] \approx \frac{1}{N}\sum_{i=1}^{N} f(\theta^{(i)})$$

где $\theta^{(i)} \sim p(\theta)$

monte_carlo_pi.png

Markov Chains

Свойство Маркова

$$P(X_{t+1} = x | X_t, X_{t-1}, \ldots, X_0) = P(X_{t+1} = x | X_t)$$

Стационарное распределение

При определенных условиях цепь сходится к распределению π, которое не зависит от начального состояния:

$$\lim_{t \to \infty} P(X_t = x) = \pi(x)$$

MCMC

Основная идея MCMC

Мы создаем марковскую цепь, чьё **стационарное распределение** совпадает с нашим целевым распределением $p(\theta)$, из которого мы хотим сэмплировать.

mcmc_concept.png

Алгоритм Метрополиса-Гастингса

Алгоритм

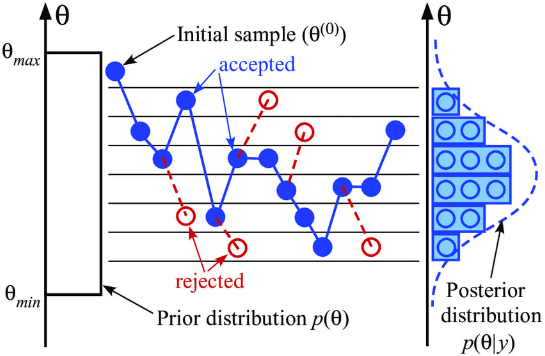
1. Инициализируем θ_0, берем распределение $q(\theta)$ из которого мы умеем сэмплировать
2. Для каждого $t = 1, 2, \ldots, N$:
   1. Предлагаем новый кандидат: $\theta^* \sim q(\theta^*|\theta_{t-1})$
   2. Вычисляем вероятность принятия:

      $$\alpha = \min\left(1, \frac{p(\theta^*) \cdot q(\theta_{t-1}|\theta^*)}{p(\theta_{t-1}) \cdot q(\theta^*|\theta_{t-1})}\right)$$

   3. Принимаем или отвергаем:

      $$\theta_t = \begin{cases} \theta^* & \text{с вероятностью } \alpha \\ \theta_{t-1} & \text{с вероятностью } 1 - \alpha \end{cases}$$

Burn-in и сходимость



Burn-in период

Первые B итераций отбрасываются, так как цепь еще не сошлась к стационарному распределению.

Популярные варианты MCMC

Hamiltonian Monte Carlo (HMC)

Использует информацию о градиенте для более эффективного исследования пространства параметров.

No-U-Turn Sampler (NUTS)

Автоматическая настройка параметров HMC. Используется в Stan.

Gibbs Sampling

Особый случай MH, где все предложения принимаются с вероятностью 1. Работает хорошо, когда можно сэмплировать из условных распределений.

Преимущества и ограничения

Преимущества

- Работает со сложными моделями
- Гибкость в постановке задачи
- Дает полное распределение, а не точечную оценку
- Относительно простая реализация

Ограничения

- Вычислительно затратно
- Сложность диагностики сходимости
- Чувствительность к настройкам
- Может плохо работать в очень больших размерностях

Выводы

Ключевые моменты

- MCMC объединяет силу Monte Carlo (усреднение по выборке) и Markov Chains (сходимость к стационарному распределению)
- Алгоритм Метрополиса-Гастингса — основа многих современных методов
  item MCMC революционизировал байесовскую статистику и прикладное моделирование
- Современные методы (HMC, NUTS) делают MCMC еще более мощным инструментом