

Lect 16 Metropolis algorithm

一. 引言:

Metropolis 算法是二十世纪十大算法之一, 在统计物理, 计算统计等领域有根本性地位. 其目标是计算

$$\langle A \rangle \triangleq \frac{1}{Z} \int_{\mathbb{R}^n} A(x) \cdot e^{-\beta H(x)} dx, \quad n \gg 1$$

其中 $Z = \int_{\mathbb{R}^n} e^{-\beta H(x)} dx$, $H(x)$ 为能量, $\beta = (k_B T)^{-1}$.

二. Metropolis 算法:

1. 直观:

由传统 MC 方法, $\langle A \rangle$ 可由

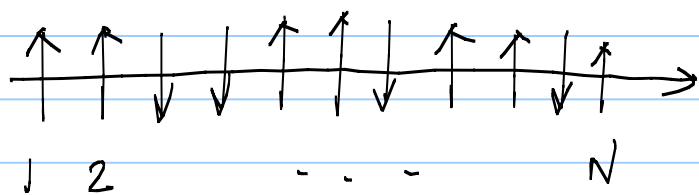
$$\langle A \rangle \approx \langle A \rangle_N \triangleq \frac{1}{N} \sum_{i=1}^N A(x_i), \quad x_i \stackrel{\text{i.i.d.}}{\sim} \frac{1}{Z} e^{-\beta H(x)}$$

逼近, 但 x_i 对应的分布难以生成!

2. Ising 模型: (1D)

考察形如

16-02



的 1D Ising 模型, 共有 N 个位点 (site), 每个位点上有一个自旋

$$\sigma_i = \begin{cases} +1 & \text{"}\uparrow\text{"} \\ -1 & \text{"}\downarrow\text{"} \end{cases}$$

记 $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_N) \in \{-1, +1\}^N$ 为系统的一个微观态。
Ising 模型为研究铁磁体相变的一个典型模型。定义微观态对应的能量

$$H(\sigma) = -J \sum_{\langle i, j \rangle} \sigma_i \sigma_j \quad J > 0 \text{ 的常数}$$

$\langle i, j \rangle$ 指最近邻相互作用, 即 $|i - j| = 1$.

统计物理中感兴趣的量为

内能

$$U_N = \frac{1}{N} \langle H \rangle = \frac{1}{N} \cdot \frac{1}{Z} \sum_{\sigma} H(\sigma) e^{-\beta H(\sigma)}$$

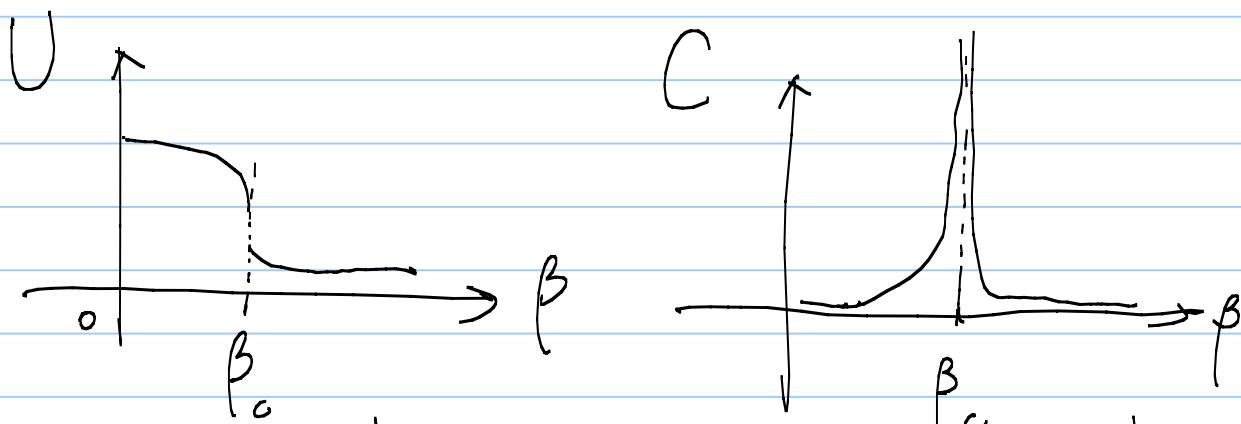
$$Z = \sum_{\sigma} e^{-\beta H(\sigma)}, \quad \beta = (k_B T)^{-1}$$

比热,

$$C_N = \frac{1}{N} V_{\text{av}}(H) = \frac{1}{N} (\langle H^2 \rangle - \langle H \rangle^2)$$

16-03

在2D时随 β 的增加, 当 $N \rightarrow \infty$ 时有如曲线:



在任意有限 N 时曲线为光滑。通过理论及数值计算寻找 β_c 及观察其行为是感兴趣的目标。

3. Metropolis算法:

由于直接生成 $\pi(\sigma) \triangleq \frac{1}{Z} e^{-\beta H(\sigma)}$ 较困难, Metropolis算法利用 Markov 链来动态生成 $\pi(\sigma)$.

设初分布为 μ_0 , 转移概率矩阵为 P , 则 n 步后

$$\mu_n = \mu_0 \cdot P^n$$

目标是让 $\mu_\infty = \pi$, 即 π 为该 Markov 链的不变分布. 由此得转移概率矩阵 P 满足

$$\pi(\sigma) = \sum_{\sigma'} \pi(\sigma') P(\sigma' \rightarrow \sigma)$$

16-04

由此不足以确定 P 矩阵.

进一步添加细致平衡条件:

$$\pi(\sigma) P(\sigma \rightarrow \sigma') = \pi(\sigma') P(\sigma' \rightarrow \sigma)$$

由此得

$$\frac{P(\sigma \rightarrow \sigma')}{P(\sigma' \rightarrow \sigma)} = \frac{\pi(\sigma')}{\pi(\sigma)} = e^{-\beta \Delta H_{\sigma \rightarrow \sigma'}}$$

$$\Delta H = H(\sigma') - H(\sigma)$$

定义 Metropolis 策略:

$$A(\sigma \rightarrow \sigma') = \min(1, e^{-\beta \Delta H_{\sigma \rightarrow \sigma'}})$$

易证

$$\frac{A(\sigma \rightarrow \sigma')}{A(\sigma' \rightarrow \sigma)} = \frac{\pi(\sigma')}{\pi(\sigma)}.$$

实际构造中取

$$P(\sigma \rightarrow \sigma') = Q(\sigma \rightarrow \sigma') A(\sigma \rightarrow \sigma'), \quad \sigma \neq \sigma'$$

其中 Q 为一转移概率矩阵, 称予选矩阵 (proposal matrix), 通常取 Q 为对称予选, 则

16-05

$$\frac{P(\sigma \rightarrow \sigma')}{P(\sigma' \rightarrow \sigma)} = \frac{A(\sigma \rightarrow \sigma')}{A(\sigma' \rightarrow \sigma)} = \frac{\pi(\sigma')}{\pi(\sigma)}$$

当 Q 为非对称情形, 则对应于 Metropolis-Hastings 算法.

当 $\sigma = \sigma'$ 时, $P(\sigma \rightarrow \sigma) \triangleq 1 - \sum_{\tilde{\sigma}} P(\sigma \rightarrow \tilde{\sigma})$

易证 P 为一转移概率矩阵.

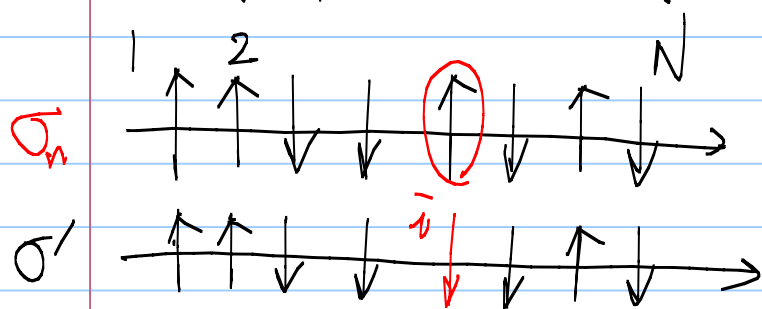
• 通常的预选方法:

a) 若可能预选: (不实际)

设当前状态 σ_n , 取 $Q(\sigma \rightarrow \sigma') = \frac{1}{2^N - 1}$

$\forall \sigma' \neq \sigma$, 而 $Q(\sigma \rightarrow \sigma) = 0$, 此取法满足条件但不切实际!

b) 单点反转预选 (Gibbs 抽样)



设当前状态 σ_n , 在 1 到 N 个自旋数中均匀随机抽取, 如左图, 设为 $\bar{i}$, 将 $\sigma_{\bar{i}}$ 的

16-06

值反转, 易证此构造方法为对称.

Metropolis算法: 设当前状态为 σ_n

Step 1: 由相应预选规则由 σ_n 产生 σ'

Step 2: 令 $\Delta H = H(\sigma') - H(\sigma_n)$, 计算

$$A(\sigma_n \rightarrow \sigma') = \min(1, e^{-\beta \Delta H})$$

Step 3: 生成 $R \sim \mathcal{U}[0, 1]$

Step 4: 如 $R < A(\sigma_n \rightarrow \sigma')$, 令 $\sigma_{n+1} = \sigma'$

否则, 令 $\sigma_{n+1} = \sigma_n$.

定理: 如马尔可夫链转移概率矩阵 P 为本原 ($\exists r \in \mathbb{N}$, s.t. $P^r > 0$ 任一分量), 则对任意初分布 ν

$$\|\nu P^n - \pi\| \rightarrow 0 \quad n \rightarrow \infty$$

其中 $\|\cdot\|$ 为全变差, $\|\mu\| \triangleq \sum_{\sigma} |\mu(\sigma)|$.

进一步结果可参考

G. Winkler, Image analysis, random fields and
dynamical Monte Carlo method, Springer.