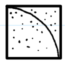
Markov Chain Monte Carlo Method

2021年8月26日 20:14

P75 Sampling Method Introduction

Mento Carlo Method : Repeaded random sampling to obtain numerical result.

e.g. $P(z|x)$ observed data, latent data   e.g. Find π



Goal : Find $E_{z|x}[f(z)] = \int p(z|x) f(z) dz$

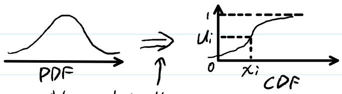
Sample $z_1 \cdots z_N \sim p(z|x)$ $E_{z|x}[f(z)] \approx \frac{1}{N}\sum_{i=1}^{N} f(z_i)$

① Probability Distribution Sampling

Random numbers between 0, 1 are eaily obtained

Get CDF (cumulative distribution function) from PDF, and then sample from the CDF according to the random numbers above

e.g.



Always difficult, even impossible

i: $u_i \sim U(0,1)$ ii: $x_i = CDF^{-1}(u_i)$ Repeat N times

$\therefore x_1 \cdots x_N \sim PDF$

② Rejection Sampling

It's difficult to sample from $p(z)$ direactly, so setting a proposal distribution $q(z)$ s.t. $\forall z_i$, $Mq(z_i) \geq P(z_i)$ (constant term: M)

Then sample from $q(z)$, and retaining samples that subject to $p(z)$

Set α : acceptance probability $\alpha = \frac{①}{②} = \frac{P(z_i)}{Mq(z_i)}$, $0 \leq \alpha \leq 1$

i: Sample $z_i \sim q(z)$ ii: Sample $u \sim U(0,1)$ iii: If $u \leq \alpha$, accept z_i, else, reject z_i Repeat N times

③ Importance Sampling

Sample from expectation : $E_{p(z)}[f(z)] = \int p(z) f(z) dz = \int f(z) \frac{p(z)}{q(z)} q(z) dz$

$\approx \frac{1}{N}\sum_{i=1}^{N} f(z_i) \frac{p(z_i)}{q(z_i)} \rightarrow$ weight $z_i \sim q(z)$, $i = 1 \cdots N$

Sampling - Importance - Resampling ?

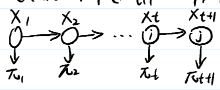
P76 Markov Chain

CDF is difficult to get and the gap between $q(z)$, $p(z)$ is large in high dimension   ∴ Use Markov Chain

Markov Chain : A stochastic model describing a sequence of possible event in which the probability of each event depends only on the state attained in the previous event (Markov Property ; memorylessness)

A sequence of possible events : $\{X_t\}$ First order Markov Chain : $P(X_{t+1} = x | X_1 \cdots X_t) = P(X_{t+1} = x | X_t)$

P : transition matrix $[P_{ij}]$ $P_{ij} = P(X_{t+1} = j | X_t = i)$



$\begin{cases} \pi = [\pi_1, \pi_2 \cdots \pi_n \cdots] \\ \sum_{i=1}^{\infty} \pi_i = 1 \end{cases}$

Stationary Distribution : $\pi(x^*) = \int \pi(x) \cdot \underbrace{P(x \mapsto x^*)}_{P(x^*|x)} dx$

e.g.

$\therefore \pi(x^*) = \sum_i \pi(x_i) \cdot P_{i*}$

∴ $\{\pi(k)\}$ is stationary distribution of $\{x_i\}$ now

Goal : By constructing a Markov Chain, so that the stationary distribution $\{\pi(k)\}$ approximates $P(z|x)$

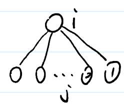
Sufficient condition for stationary distribution : Detailed Balance : $\pi(A) P(A \mapsto B) = \pi(B) P(B \mapsto A)$

Detailed Balance $\Rightarrow$ Stationary Distribution

$\int \pi(x) p(x \mapsto x^*) dx = \int \pi(x^*) p(x^* \mapsto x) dx^* = \pi(x^*) \underbrace{\int p(x^* \mapsto x) dx^*}_{1} = \pi(x^*)$

A row of the trasition matrix

$\sum_{j=1}^{\infty} P_{ij} = 1$



P77 Metropolis-Hastings Algorithm

Goal : Find a transition matrix $P = [P_{ij}]$ subject to detailed balance

Set a random transition matrix $Q = [Q_{ij}]$ and a factor $\alpha(i,j)$

$P(A) \underbrace{Q(A \mapsto B) \alpha(A,B)}_{P(A \mapsto B)} = P(B) Q(B \mapsto A) \alpha(B,A)$

$\alpha(A,B) = \min\left(1, \frac{P(B)Q(B \mapsto A)}{P(A)Q(A \mapsto B)}\right) \longleftarrow$ weight

Proof. : Left $= P(A)Q(A \mapsto B)\alpha(A,B) = P(A)Q(A \mapsto B) \min\left(1, \frac{P(B)Q(B \mapsto A)}{P(A)Q(A \mapsto B)}\right)$

$= \min\left(P(A)Q(A \mapsto B), P(B)Q(B \mapsto A)\right) = P(B)Q(B \mapsto A) \min\left(1, \frac{P(A)P(A \mapsto B)}{P(B)P(B \mapsto A)}\right)$

$= P(B)Q(B \mapsto A)\alpha(B,A) = $ Right

Sampling : $u = U(0,1)$ $z^* \sim Q(z|z_{i-1})$? $\alpha = \min\left(1, \frac{P(z^*)Q(z^* \mapsto z)}{P(z)Q(z \mapsto z^*)}\right)$

If $u \leq \alpha$, $z_i = z^*$ else $z_i = z_{i-1}$

$= P(B)Q(B \mapsto A)\alpha(B,A) =$ Right

Sampling: $u = U(0,1)$ $\quad z^* \sim Q(z | z_{i-1})$? $\quad \alpha = \min\left(1, \frac{P(z^*)Q(z^* \mapsto z)}{P(z)Q(z \mapsto z^*)}\right)$

If $u \le \alpha$, $z_i = z^*$, else $z_i = z_{i-1}$