Welcome

resume at 2:02!

COMP 136

Day 20: Sampling Methods & Intro to MCMC

sorry posted late due to technical difficulties

part 1 + part 2

Agenda:

1:30 – 2:00 : extra time to watch videos

2:00 – 2:05 : announcements

2:05 – 2:45 : exercises/Q&A

if you have watched all videos, ask questions on Piazza!

gradescope

Upcoming Due Dates

- Quiz3 released, due **Thu Apr 9** 11:59pm ET

---- Will be entirely online via gradescope

---- Time limit: 20 minutes

---- all multiple choice

Unit 3

Kmeans

Gaussian mixtures

HW4 released! (due next Wed at 11:59pm ET)

CP4 released in next few days (stay tuned for due date)

New Idea #1: Monte Carlo Estimation

Rand var: z w/ pdf $p(z)$

Expectation:

$$\bar{f} = \mathbb{E}_{z \sim p(z)}[f(z)] = \int f(z) p(z) dz$$

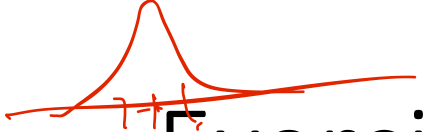
MC estimate

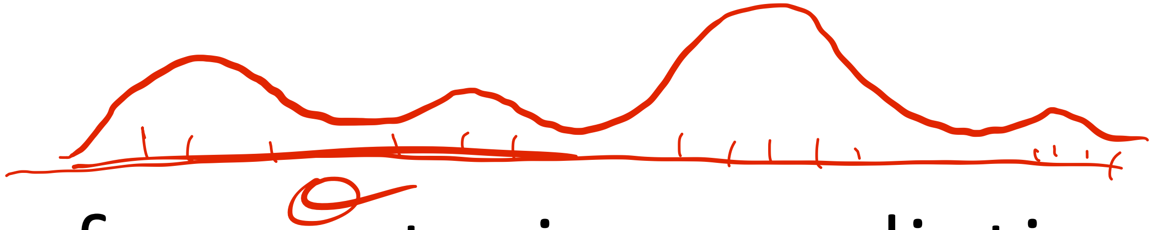
$$\hat{f} = \frac{1}{S} \sum_{s=1}^{S} f(z^s)$$

S samples of $p(z)$

$z^1, z^2, \ldots z^S$

unbiased $\mathbb{E}[\hat{f}] = \bar{f}$, variance $\frac{1}{S}$





Exercise: MC estimator for posterior predictive

$p(\theta)$ prior

$p(x_n|\theta)$ likelihood

$\{x_n\}_{n=1}^N$ train

x_* test

- Suppose you can sample from posterior

$$\theta_1, \theta_2, \cdots \theta_S \sim p(\cdot \mid x_{1:N})$$

tractable all Gaussian regression

- Show how you can estimate the posterior predictive likelihood

Sum rule

$$p(x_* \mid x_{1:N}) = \boxed{\int p(x_*|\theta)\, p(\theta|x_{1:N})\, d\theta}$$

$$= \mathbb{E}_{p(\theta|x_{1:N})}\left[p(x_*|\theta) \right]$$

$$= \frac{1}{S} \sum_{s=1}^{S} p(x_*|\theta^s)$$

$$\hat{f} = \frac{1}{S} \sum_{s=1}^{S} f(z^s)$$

MC estimator

r.v. $z \leftarrow \theta$

pdf $p(z)$ $p(\theta|x_{1:N})$

f $f(z)$ $p(x_*|\theta)$

works even for high dim θ

New Idea #2: Ancestral Sampling

Joint distribution = product of simple conditionals

$$p(z_1) \cdot p(z_2 \mid z_1) \cdot p(z_3 \mid z_2, z_1) \cdot p(z_4 \mid z_3, z_2, z_1)$$

order node indices

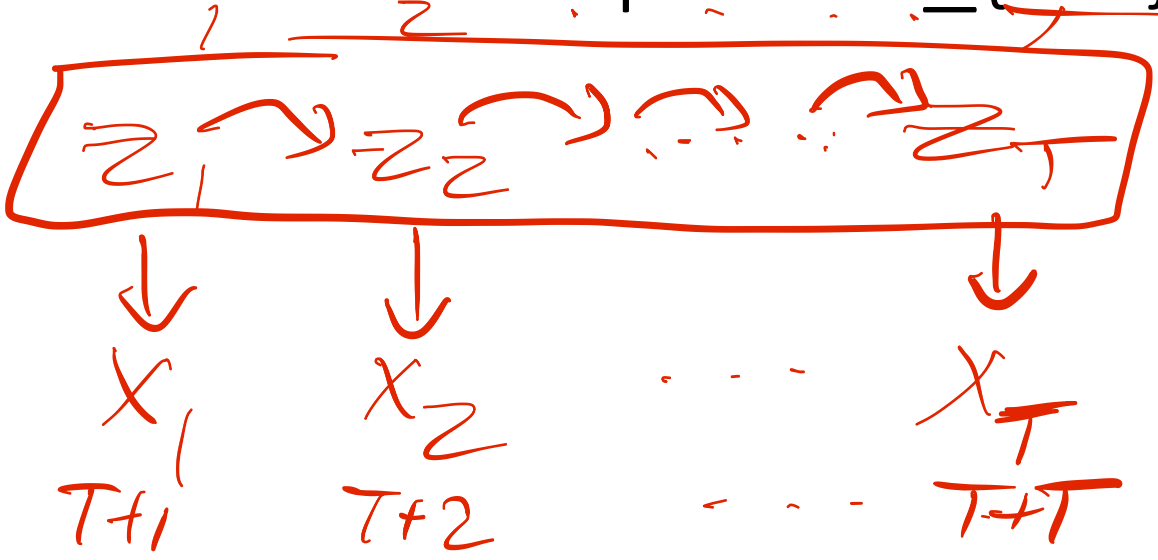
s.t. i after its parents

Exercise:
How to do Ancestral Sampling for HMM?

$A = \begin{bmatrix} \text{—} \\ \text{—} \\ \vdots \\ \text{—} \end{bmatrix}$

Given: HMM parameters $\underbrace{\text{pi, A, mu, sigma}}$ $= \theta$

Want: sample of z_{1:T} and x_{1:T} from $p(x_{1:T}, z_{1:T} \mid \theta)$



$x_t \sim p(x_t \mid z_t = k, \theta_k)$

1	$z_1 \sim \mathrm{Cat}(\pi)$	$x_1 \sim N(\mu_{z_1}, \sigma^2_{z_1})$
2	$z_2 \mid z_1 \sim \mathrm{Cat}(A_{z_1})$	$x_2 \sim N(\mu_{z_2}, \sigma^2_{z_2})$
$\vdots$	$\vdots$	$\vdots$
T	$z_T \mid z_{T-1} \sim \mathrm{Cat}(A_{z_{T-1}})$	$x_T \sim N(\mu_{z_T}, \sigma^2_{z_T})$
	$z_{1:T} \mid \theta$	$x_{1:T} \mid z, \theta$

New Idea #3: Markov Chain Monte Carlo

p^* is a valid pdf $c = \frac{1}{\int \tilde{p}(z) dz}$

day 16 notebook in github!

Target: p^* over r.v. z (cont / discrete) z_k

stationary distrib.

p^* difficult to sample from

design transition proposal

$p^*(z) = c\,\tilde{p}(z)$

unknown constant $c > 0$ known func $\tilde{p}(z) \geq 0$

across many samples

	$p(z_S)$	p^*
$z_S = 1$	0.25	
$z_S = 2$	0.50	
⋮	0.10	
$z_S = 4$	0.15	

prev z value → samples z values

$z_1 \to z_2 \to z_3 \to z_4 \cdots z_S$

init value

$S >> 0$ 1000

$p(z_S) = p^*(z)$

p^* is stationary distrib. of Markov chain

New Idea #4: Inverse CDF Sampling Method

Consider for example the *exponential distribution*

$$p(y) = \lambda \exp(-\lambda y) \tag{11.7}$$

where $0 \leqslant y < \infty$. In this case the lower limit of the integral in (11.6) is 0, and so $h(y) = 1 - \exp(-\lambda y)$. Thus, if we transform our uniformly distributed variable z using $y = -\lambda^{-1} \ln(1 - z)$, then y will have an exponential distribution.

Use CDF inverse method

$$F(y) = 1 - e^{-\lambda y} = u$$

So $F^{-1}(u) = -\frac{1}{\lambda} \log(1-u)$ by algebra

To Sample y:

1) Draw $u \sim$ Unif over $[0,1]$

2) Set $y \leftarrow F^{-1}(u)$

New Idea #5: Transforms of Random Vars

- Given u ~ Normal(0, 1)
- Transform: x ← T(u) = s * u + m, where s > 0 and m is a real value
- What is the PDF of x?

$$f(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

$$x = su + m$$

$$\frac{x-m}{s} = u \qquad S(x) = \frac{x-m}{s} \qquad S'(x) = \frac{1}{s}$$

$$\begin{tikzcd} u \arrow[r, bend left, "S"] & x \arrow[l, bend left, "T"] \end{tikzcd}$$

abs val

$$\text{pdf}(x) = f(S(x))\,|S'(x)|$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-m)^2}{s^2}} \frac{1}{s}$$

assumes $s > 0$ so abs val goes away

$$= \text{NormPDF}(x \mid m, s^2)$$