Gibbs Sampling

SPR Day 22

IN CLASS

Agenda

HW4 help Q&A
CP4 help Q&A
Q&A on Gibbs

Coming Up

HW4 due tonight
CP4 due Tues
Q4 late next week

Mon		Wed
		15
~~class~~ OH	CP4	class
class		

HW5 out this week
CP5 out this week
due Mon 5/4

HW4

1 proove conditional indep using sum/product rules

2 EM alg. for HMMs

3 stationary of Markov

computational / analytical [0.--- ___ ___ ___] non-neg. sum to one

defines **PMF** of discrete r. variable

Hint: See Notebook

1b asks:

assumed HMM $0:T$

$p(x_{1:T}, z_{1:T})$

$$\boxed{p(x_{t+1} \mid x_{1:t}, z_{1:t})} = p(x_{t+1} \mid z_t)$$

Hint

(1) $p(x_{t+1} \mid z_{1:t})$

Bayes/product rule

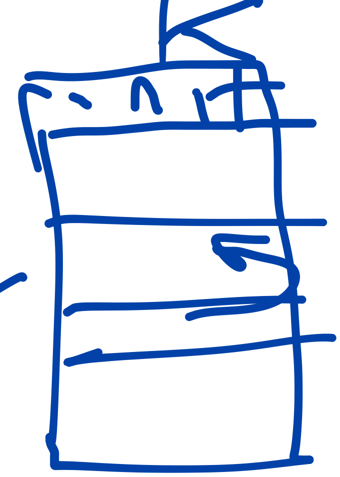
(2) $p(x_{t+1} \mid z_t)$ z_{t+1}

$\boxed{\text{Sum rule}}$

~~prod rule~~

CP4

forward



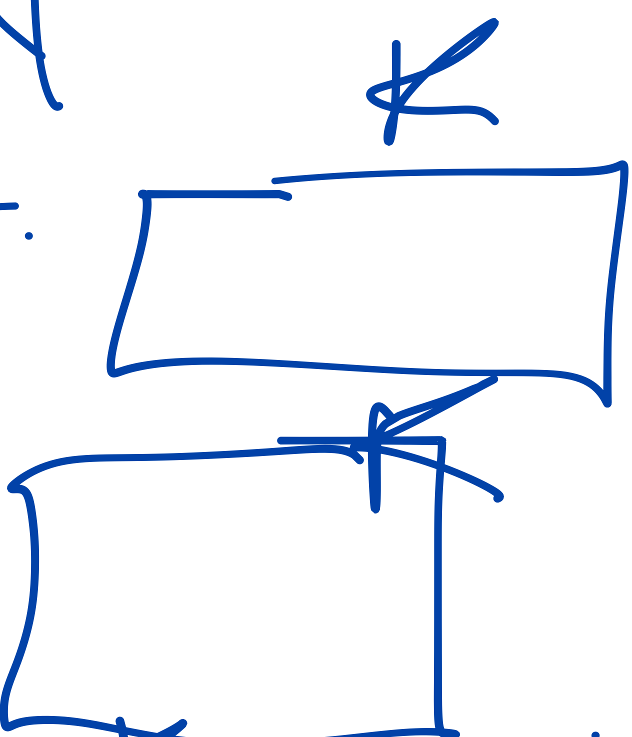
$a_{t+1} = f(a_t)$

viterbi

---

input

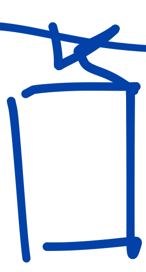
T



sums to 1

A K

L T



log lik

$\log w_{tk}$

log joint prob of most likely seq that ends at state k at t

b_{tk}

Gibbs

$z_1 \cdots z_D \sim p^*$

Assume: joint HARD

Cond. EASY

$z_1 \cdots z_D$

$z_1 \sim p^*(z_1 | z_2, \ldots, z_D)$

$\boxed{z_1}$ $z_2 \cdots z_D$

z_1 $z_2 \cdots z_D$

z_1 $z_2 \cdots z_D$

$z_1^{t+1} \cdots z_D^{t+1}$

GMM

$x_n \in \mathbb{R}$ (1 dim features)

RV: π, μ, τ, x, z

Model:

$\pi \sim \mathrm{Dir}(\alpha)$ ✓

$\mu_k \sim N(0, \beta^{-1} I)$ ✓

$\tau_k \sim \mathrm{Gam}(a, b)$ ✓

$z_n \sim \mathrm{Cat}(\pi_1 \cdots \pi_K)$ ✓

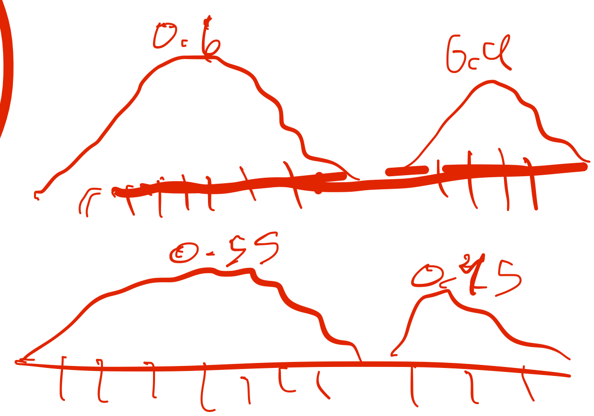
$x_n | z_n = k \sim N(\mu_k, \tau_k^2)$ ✓

ancestral sampling of full joint ✓

Target: Posterior given Data

$$\underline{p(\pi, \mu, \sigma, z} \mid x)$$

K=2



this itself is a "joint" in sense that π, μ, σ, z together form a big vector

Easier conditional:

$$p(\boxed{z_n} \mid \underline{\overset{x_n}{\cancel{x_{1:N}}}, \cancel{z_{\setminus n}}, \pi, \mu, \sigma}) \quad ?$$

discrete K possible values $\{1, 2, \ldots K\}$

(1) What is sample space?

(2) what <u>indep properties</u>?

(3) <u>Similar to from EM</u>?

GMM assumptions

$$p(z_{1:N}|\pi) = \prod_n p(z_n|\pi)$$

z_n indep of z_i, $i \neq n$, given π

$$p(x_{1:N}|z_{1:N}, \cancel{\pi}, \mu, \sigma) = \prod_n \prod_k \mathcal{N}(x_n|\mu_k, \sigma_k^2)^{\delta(z_n,k)}$$

x_n indep of z_i given z_n

$$p(z_n | \overbrace{x_{1:N}, z_{\backslash n}, \pi, \mu, \sigma}^{\text{others}})$$

$$\propto p(z_n|\pi)\, p(x_n|\mu_{z_n}, \sigma_{z_n})$$

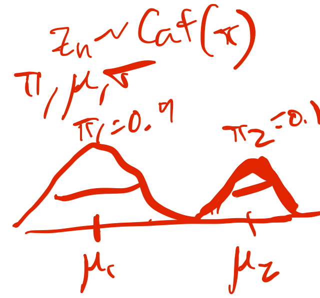
$$\underbrace{p(z_n = k | x_n, \pi, \mu, \sigma)}_{\gamma_{nk}} = \boxed{\frac{\pi_k \,\text{NormPDF}(x_n|\mu_k, \sigma_k^2)}{\sum_{l=1}^{K} \pi_l \,\text{NormPDF}(x_n|\mu_l, \sigma_l^2)}}$$

Conditional:

$$\boxed{p(z_n|\text{others}) = \text{Cat}(\gamma_{n1}, \ldots, \gamma_{nK})}$$

draw sample from categorical distrib. w/ params γ_n

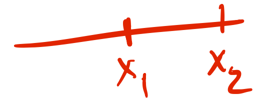
Gibbs for GMM



$z_n \sim Cat(\pi)$

$z_1 = 1$

$z_2 = 2$



Initialize $z_{1:N}, \pi, \mu, \sigma$

for t in $2, 3, \ldots S$:

for each n:

$z_n^t \sim p(z_n \mid \text{others})$

$\pi^t \sim p(\pi \mid z_{1:N}, \mu, \sigma, x_{1:N})$

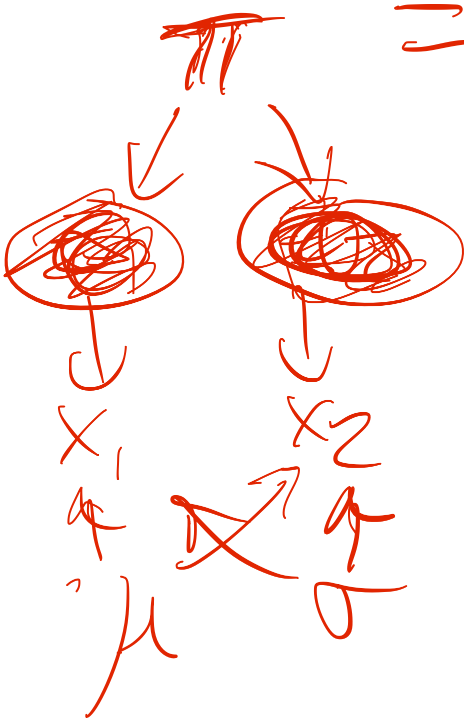
$\mu^t \sim p(\mu \mid z_{1:N}, \pi, \sigma, x_{1:N})$

$\sigma^t \sim p(\sigma \mid \text{others})$

Gibbs is an MCMC method

Example of simplifying conditional

$$p(\pi \mid x_{1:N}, \underline{z_{1:N}}, \mu, \sigma)$$



$$= p(\pi \mid z_{1:N})$$ by indep assumptions

$$= \frac{p(\pi)\, p(z_{1:N} \mid \pi)}{p(z_{1:N})}$$ Bayes

$$= \mathrm{Dir}(\pi \mid \alpha) \prod_{n=1}^{N} \mathrm{Cat}(z_n \mid \pi)$$

conjugacy of Dir + Cat CP1

$$= \mathrm{Dir}(\pi \mid N + \alpha)$$

$$= \mathrm{Dir}(N_1 + \alpha_1, \ldots N_K + \alpha_K)$$

$N_K = \# n$ s.t. $z_n = K$

sample

copy

after t: z^t π^t μ^t σ^t

$t+1$ z: z^{t+1} π^t μ^t σ^t — concrete values — $p^*(z \mid \pi^t, \mu^t, \sigma^t, x_{1:N})$

$t+1$ π: z^{t+1} π^{t+1} μ^t σ^t — $p^*(\pi \mid z^{t+1}, \mu^t, \sigma^t, x_{1:N})$

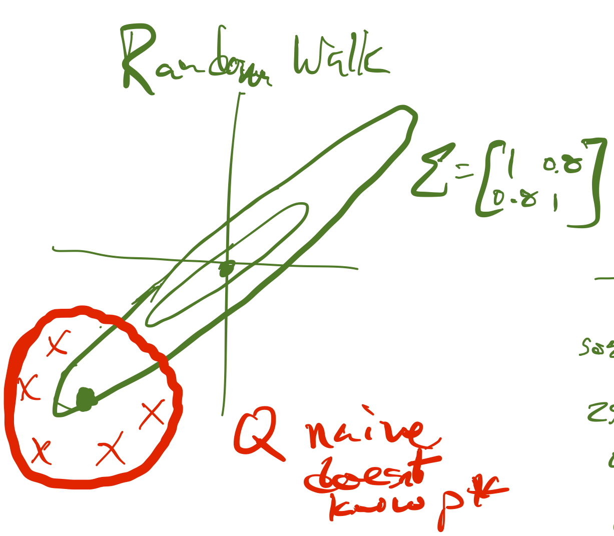
$t+1$ μ: z^{t+1} π^{t+1} μ^{t+1} σ^t — $p^*(\mu \mid \pi^{t+1}, z^{t+1}, \sigma^t, x_{1:N})$

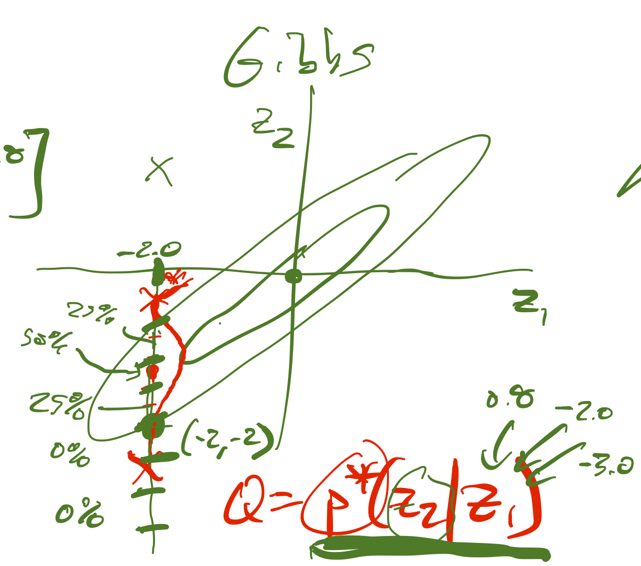
$t+1$ σ: z^{t+1} π^{t+1} μ^{t+1} σ^{t+1} — $p^*(\sigma \mid \pi^{t+1}, z^{t+1}, \mu^{t+1}, x_{1:N})$

pick order: z then π, then μ, then σ

σ then π, then μ, then z

σ then z then σ, then π, then μ





= level set contains 99% of proba mass

O = Q proposal pdf