Gibbs Sampling (continued)

$$P(x,y) = y \exp\left(-[xy + y]\right)$$

$$P(x|y) = \frac{P(x,y)}{\int dx' \, P(x', y)}$$

$$P(x|y) \propto \exp(-xy) \qquad \exp(-x)$$

Correction from Wednesday's Lecture.

Gibbs Sampling (continued)

$$p(x,y) = y \exp\left(-[xy + y]\right)$$

$$p(x|y) = \frac{p(x,y)}{\int dx\, p(x,y)}$$

$$p(x|y) \propto \exp(-xy) \qquad \exp(-x)$$

correction from Wednesday's lecture.

(simple) Gibbs, ~~☐~~ satisfies DB with $\pi = P$.

(sweep) Gibbs, <u>doesn't</u> satisfy DB.
does have P as stationary dist.

Recall $\displaystyle\int P(x)\, p(x'; x)\, dx = P(x')$

if true, P is stationary w.r.t. Markov chain $p(x'; x)$.

consider case $d=2$.

transition $x \to y$ happens in two stages $(x^0, x^1) \to (y^0, x^1) \to (y^0, y^1)$
(x^0, x^1) (y^0, y^1)

using 1-dim
conditional dists.

transition $x \to y$ happens in two stages $(x_0, x_1) \to (y_0, x_1) \to (y_0, y_1)$

(x_0, x_1) (y_0, y_1)

using 1-dim conditional dists.

consider case $d = 2$.

if true, p is stationary w.r.t. Markov chain $q(x'; x)$.

Recall $\int p(x)\, q(x'; x)\, dx = p(x')$

(sweep) Gibbs, doesn't satisfy DB.
does have p as stationary dist.

(simple) Gibbs, ~~satisfies~~ satisfies DB with $\pi = p$.

$$\rho([y^0, y^1], [x^0, x^1]) = P(y^0 | x^1)\, P(y^1 | y^0)$$

$$= \frac{P(y^0, x^1)}{\int da\, P(a, x^1)} \; \frac{P(y^0, y^1)}{\int db\, P(y^0, b)}$$

sub into stat condition.

$$\int dx^0 \int dx^1 \; P(x^0, x^1)\; \rho([y^0, y^1], [x^0, x^1]) \overset{?}{=} P(y^0, y^1)$$

$$= \int dx^0 \int dx^1 \; P(x^0, x^1)\; \underbrace{\frac{P(y^0, x^1)}{\int da\, P(a, x^1)}}\; \underbrace{\frac{P(y^0, y^1)}{\int db\, P(y^0, b)}}$$

$$= \int dx^0 \; P(x^0, x^1) \frac{P(y^0, y^1)}{\int db\, P(y^0, b)} \int dx^1 \frac{P(y^0, x^1)}{\int da\, P(a, x^1)}$$

$$= \int dx_0\, p(x_0,x_1)\, \frac{\int d\beta\, p(x_0,\beta)}{p(x_0,x_1)} \int dx_1\, \frac{\int d\alpha\, p(\alpha,x_1)}{p(x_0,x_1)}$$

$$= \int dx_0 \int dx_1\, p(x_0,x_1)\, \underbrace{\frac{\int d\alpha\, p(\alpha,x_1)}{p(x_0,x_1)}}\, \underbrace{\frac{\int d\beta\, p(x_0,\beta)}{p(x_0,x_1)}}$$

$$\int dx_0 \int dx_1\, p(x_0,x_1)\, \sigma([x_0,x_1],[x_0,x_1]) \overset{?}{=} p(x_0,x_1)$$

Sub into 2nd condition.

$$= \frac{\int d\alpha\, p(\alpha,x_1)}{p(x_0,x_1)} \frac{\int d\beta\, p(x_0,\beta)}{p(x_0,x_1)}$$

$$\sigma([x_0,x_1],[x_0,x_1]) = p(x_0|x_0)\, p(x_1|x_0)$$

$$= \frac{P(y^0, y')}{\cancel{\int dz\, P(y^0, z)}} \cancel{\int dx'} \frac{\cancel{P(y^0, x')}}{\cancel{\int da\, P(a, x')}} \cancel{\int dx^0\, P(x^0, x')}$$

$$= P(y^0, y')$$

$$= P(y) \checkmark \qquad \square$$

$= \sigma(z)$ $\lfloor$

$= \sigma(z_0, z_1)$

$\boxed{\int dz \, \sigma(z_0, z)}$

$= \sigma(z_0, z_1)$ $\boxed{\int dx_1 \, \sigma(z_0, x_1)}$ $\boxed{\int dx \, \sigma(x, x_1)}$ $\boxed{\int dx_0 \, \sigma(x_0, x_1)}$

$$[x^0, x^1] \longrightarrow [y^0, x^1] \longrightarrow [y^0, y^1]$$

$$\rho([y^0, y^1], [x^0, x^1]) = P(y^0 | x^1)\, P(y^1 | \underbrace{x^0}_{y^0})$$

this will <u>Not</u> converge to P.

this will have a stat dist $Q(x^0, x^1) = Q^0(x^0)\, Q^1(x^1)$

"ignore" correlations in target dist.

This will form a start dist $Q(x_0, x_1) = Q_0(x_0)\, Q_1(x_1)$

"ignore" correlations in target dist.

This will NOT converge to P.

$$P([x_0, x_1], [y_0, y_1]) = P(y_0 \mid x_1)\, P(y_1 \mid \underbrace{x_0}_{y_0})$$

$$[x_0, x_1] \longrightarrow [y_0, x_1] \longrightarrow [y_0, y_1]$$

DB $\quad P(x)\, e(x', x) = P(x')\, e(x, x')$

$$\Downarrow$$

stationary $\quad \int dx\, \underbrace{P(x)}_{\text{current point}}\, e(\underbrace{x'}_{\text{new point}}, x) = P(x')$

DB is (usually) easier to check because local.

Stationarity, a global property, is harder.

DB $\quad P(x)\, Q(x'|x) = P(x')\, Q(x|x')$

stationary $\quad \int dx\, P(x)\, Q(x'|x) = P(x')$

DB is (usually) easier to check because local.

Stationarity, a global property, is harder.

"Blocked" Gibbs sampler.

$$P(\underbrace{x^0, x^1}, \underbrace{x^2, x^3})$$

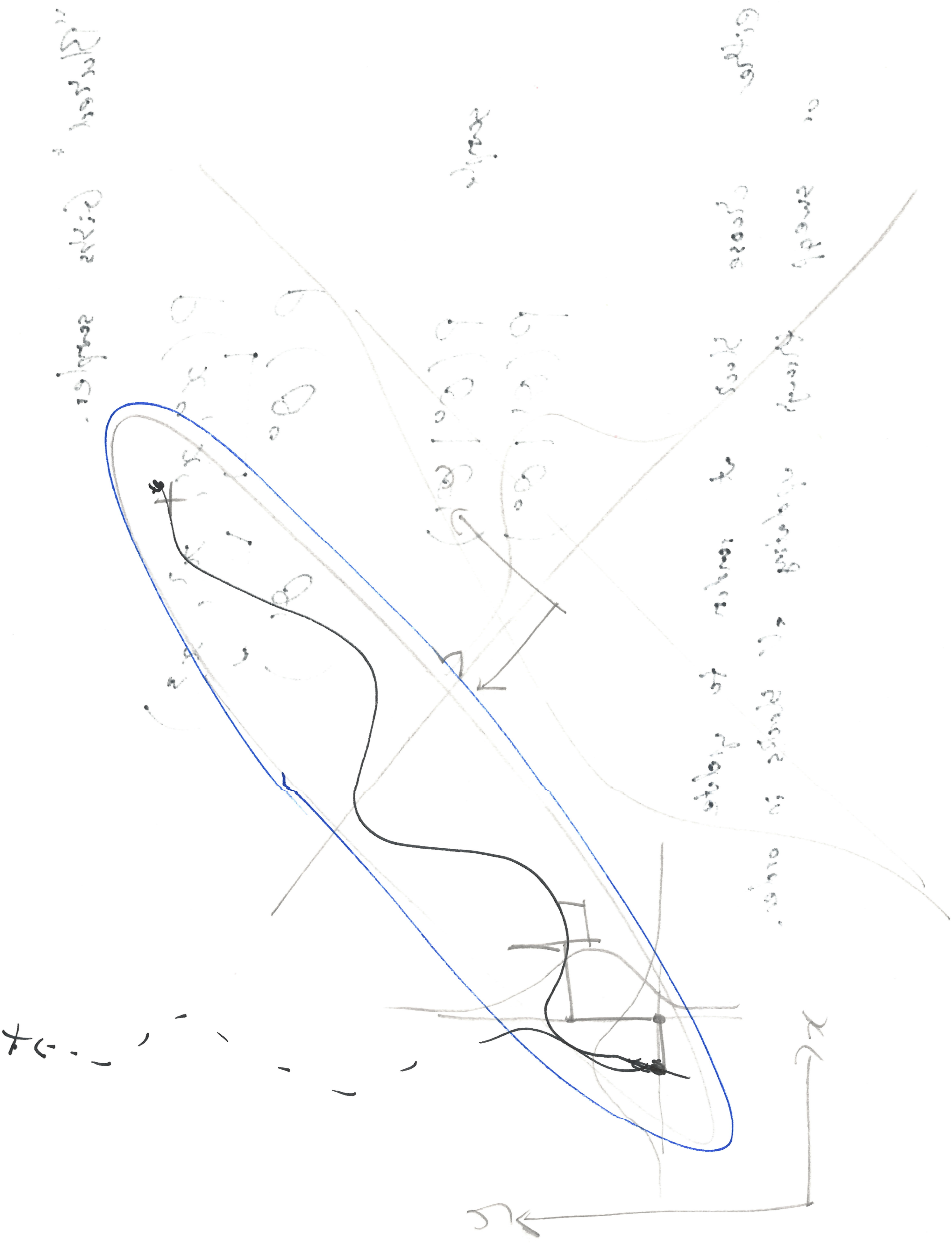
$$P(\theta^0 \quad \theta^1)$$

sample $P(\theta^0 | \theta^1)$

$P(\theta^1 | \theta^0)$

either choose block at random to update

or sweep through updating all blocks in order.



for someone familiar with MH

- have to choose Q (proposal)
- accept/reject α $\alpha = \min\left(1, \frac{p(y)}{p(x)} \frac{Q(x|y)}{Q(y|x)}\right)$

let $Q = w_k \, \delta^{(d-1)}(x^{-k} - y^{-k}) \, P(x^k | x^{-k})$

$\Rightarrow \alpha = 1$ always.

for someone familiar with Gibbs

extra freedom to use any proposal (not just conditionals)

if you reject some points

for someone familiar with MH

- have to choose Q (proposal)
- accept/reject $\alpha = \min\left(1, \frac{P(y)\, Q(x \mid y)}{P(x)\, Q(y \mid x)}\right)$

let $Q = w_k\, \delta_{(x_{-k})}(y_{-k})\, P(y_k \mid x_{-k})$

$\Rightarrow \alpha = 1$ always.

for someone familiar with Gibbs

[illegible] Gibbs to use any proposal (not just conditionals)

it can reject some points