Gibbs Sampler

- Gibbs sampler is used when samples consist of multiple componenent, e.g. $(x_1, x_2, \ldots, x_n)$.
- The joint density $f(x_1, \ldots, x_n)$ is not known (or complicated to obtain).
- But $f(x_i \mid x_1, \ldots, x_{i-1}, x_{i+1}, \ldots x_n)$ are given.

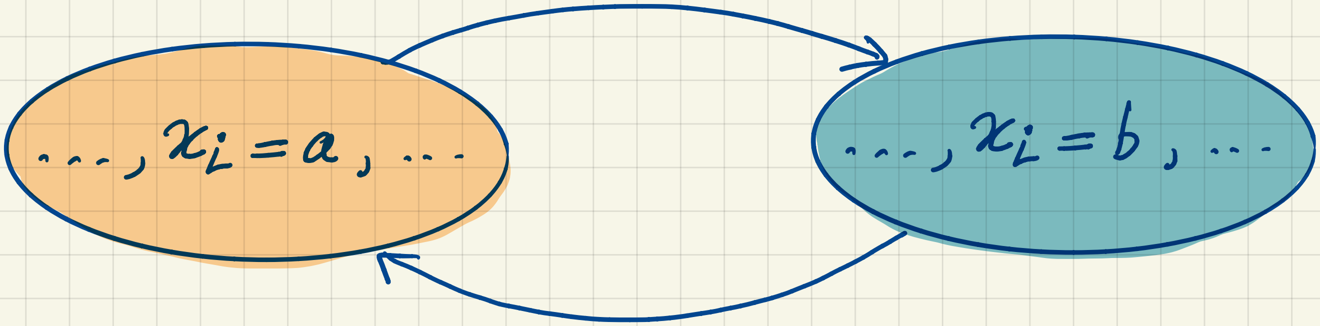
Assume $(x_1^t, \ldots x_n^t)$ is a sample at time t. Then

$$x_1^{t+1} \sim f(x_1 \mid x_2^t, \ldots, x_n^t)$$
$$x_2^{t+1} \sim f(x_2 \mid x_1^{t+1}, x_3^t, \ldots, x_n^t)$$
$$x_3^{t+1} \sim f(x_3 \mid x_1^{t+1}, x_2^{t+1}, x_4^t \ldots, x_n^t)$$
$$\vdots$$
$$x_n^{t+1} \sim f(x_n \mid x_1^{t+1}, \ldots x_{n-1}^{t+1})$$

claim: $(x_1^t, x_2^t, \ldots, x_n^t) \sim f(x_1, x_2, \ldots, x_n)$

f satisfies detailed balance "component-wise"

$$f(x_i = b \mid x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$$



$$f(x_i = a \mid x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$$

$$f(x_1, \ldots, x_i = a, \ldots x_n)\, f(x_i = b \mid x_1 \ldots x_{i-1}, x_{i+1}, \ldots, x_n)$$

$$= f(x_1, \ldots, x_i = a, \ldots x_n) \frac{f(x_1, \ldots x_{i-1}, x_i = b, x_{i+1}, \ldots x_n)}{f(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)}$$

$$= f(x_i = a \mid x_1, \ldots, x_{i-1}, x_{i+1}, \ldots x_n)\, f(x_1, \ldots, x_{i-1}, x_i = b, x_{i+1}, \ldots x_n)$$

Gibbs can be viewed as an instance of Metropolis-Hastings

* α, the prob. of accept is always 1
* the proposal distribution alternates among n distributions, one for each component.

For component i, we propose $x_i^{\delta+1} \sim f(x_i \mid x_1^{\delta+1}, \dots, x_{i-1}^{\delta+1}, x_{i+1}^{\delta}, \dots, x_n^{\delta})$

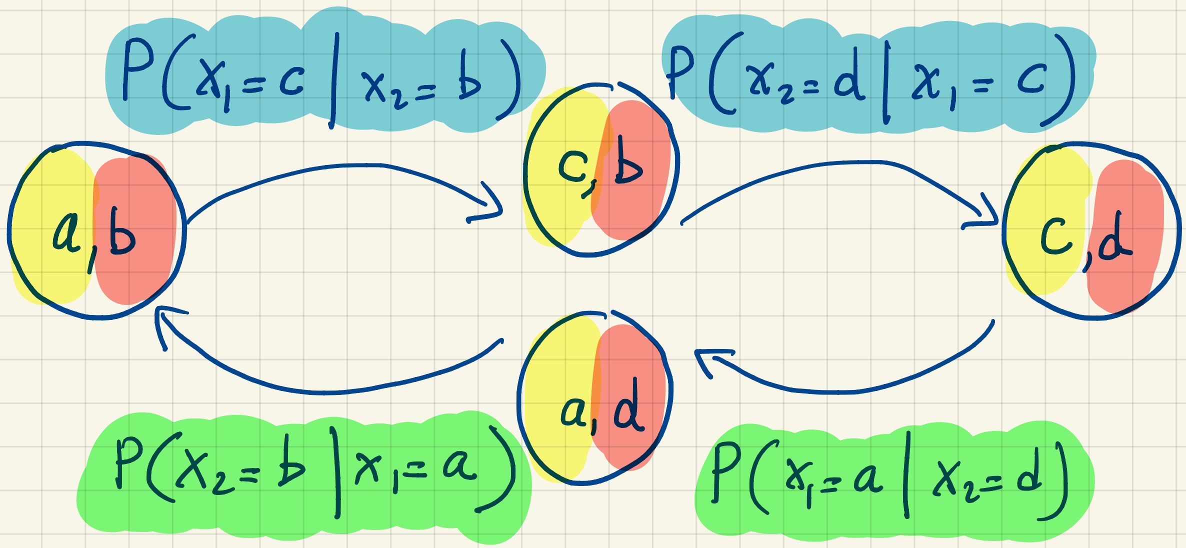
and $x_j^{\delta+1} = x_j^{\delta}$ with prob. 1 for $j \neq i$

we have $n\delta$ steps within one t step

Remark: Gibbs does NOT satisfy detailed balance for entire sample. (one big step)

e.g. $P(x_1=a, x_2=b)\, P(x_1=c, x_2=d \mid x_1=a, x_2=b)$

$\neq P(x_1=c, x_2=d)\, P(x_1=a, x_2=b \mid x_1=c, x_2=d)$



Example:

		x_2	
		0	1
x_1	0	$1/3$	$1/3$
	1	0	$1/3$

$$\underset{t}{(1,1)} \to (0,1) \to \underset{t+1}{(0,0)}$$

$$P(0,0)\,P(1,1 \mid 0,0) = P(1,1)\,P(0,0 \mid 1,1)$$

$P(0,0) = 1/3$ $\qquad$ $P(1,1) = 1/3$

$$P(1,1 \mid 0,0) = 0$$

$$(0,0) \xrightarrow{\underset{0}{P(x_1=1 \mid x_2=0)}} (1,0) \xrightarrow{\underset{1}{P(x_2=1 \mid x_1=1)}} (1,1)$$

$$P(0,0 \mid 1,1) = 1/4$$

$$(1,1) \xrightarrow{\underset{1/2}{P(x_1=0 \mid x_2=1)}} (0,1) \xrightarrow{\underset{1/2}{P(x_2=0 \mid x_1=0)}} (0,0)$$

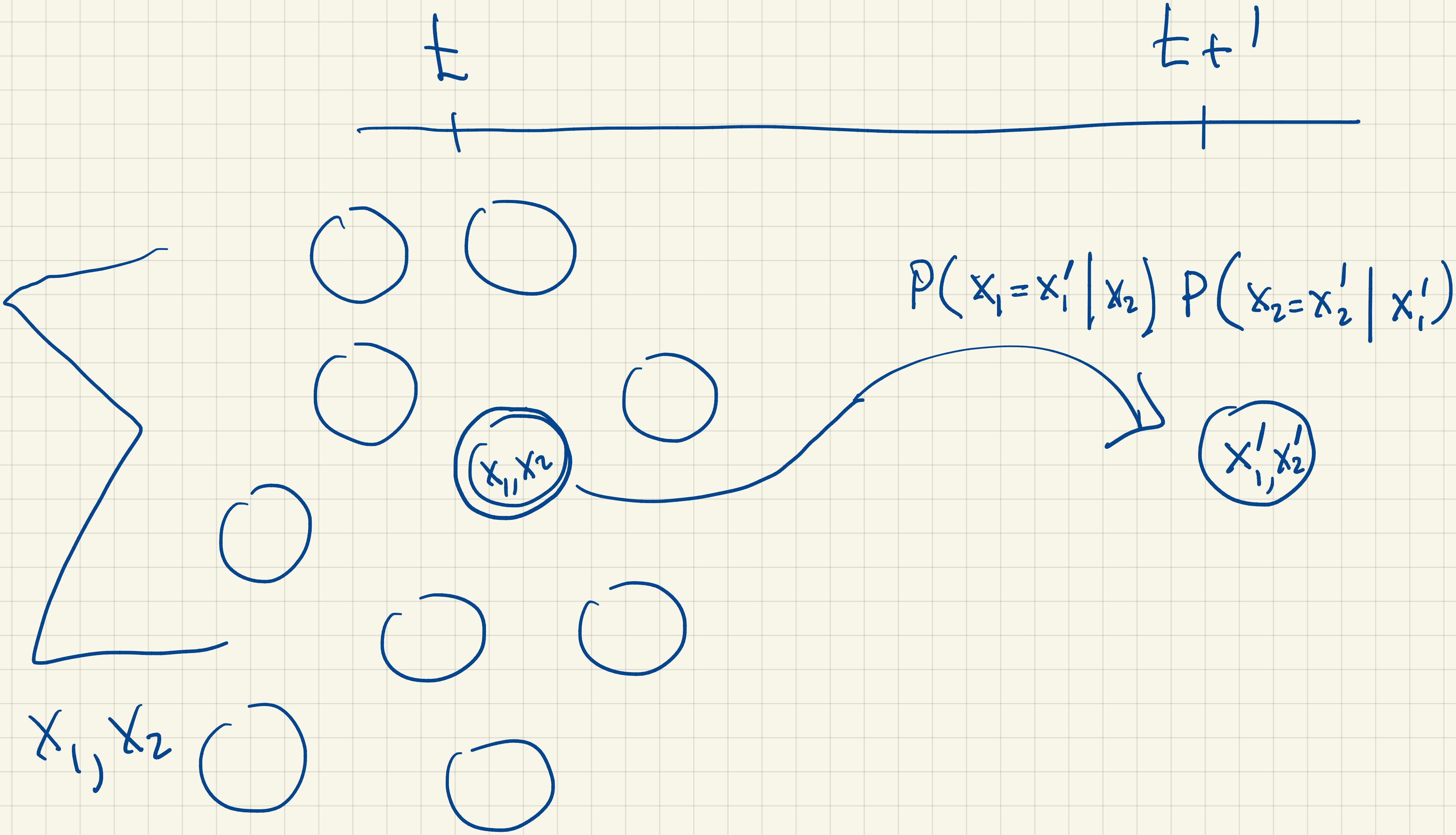
$$1/3 \times 1/4 \neq 1/3 \times 0$$

Showing that $P(x_1, x_2)$ is stationary

$$\sum_{x_1, x_2} P(x_1, x_2) \cdot P(x_1' | x_2)\, P(x_2' | x_1')$$

What's the prob. that our next sample is (x_1', x_2')

$$= \underbrace{P(x_2' | x_1') \underbrace{\sum_{x_2} \Big(\underbrace{P(x_1' | x_2) \underbrace{\sum_{x_1} P(x_1, x_2)}_{P(x_2)}}_{P(x_1', x_2)} \Big)}_{P(x_1')}}_{P(x_1', x_2')}$$



$P(x_1, x_2)$

Application to Bayes

$$x_i \mid \mu, \sigma^2 \sim N(\mu, \sigma^2)$$

priors: $\mu \sim N(\beta, \tau^2)$

$\dfrac{S_0}{\sigma^2} \sim \chi^2_k$ [semi-conjugate prior]

Conditional posteriors:

$$\mu \mid \sigma^2, \bar{x} \sim N\left(\frac{\sigma^2 \beta/n + \tau^2 \bar{x}}{\sigma^2/n + \tau^2}, \frac{\sigma^2 \tau^2/n}{\sigma^2/n + \tau^2} \right)$$

$$\frac{S+S_0}{\sigma^2} \,\Big|\, \mu, \bar{x} \sim \chi^2_{k+n} \quad \text{where } S = \sum (x_i - \mu)^2$$

Gibbs: start with some (μ^0, σ^0)

repeat $t = 0, 1, 2, 3, \ldots$

sample $\mu^{t+1} \sim N(\sim\!\sim, \sim\!\sim)$ given σ^t

sample σ^{t+1} using $\left(\dfrac{S+S_0}{\sigma^2}\right)^{t+1} \sim \chi^2_{k+n}$ given μ^{t+1}

$f(\mu, \sigma \mid \bar{x})$ is stationary for this Gibbs.