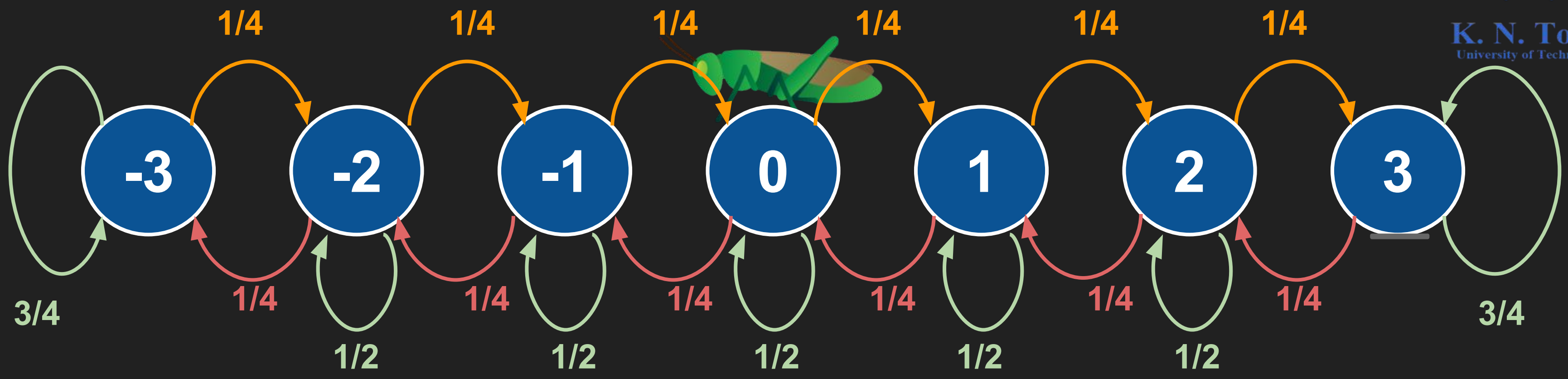
Probabilistic Graphical Models

Lectures 21

Markov Chain Monte Carlo

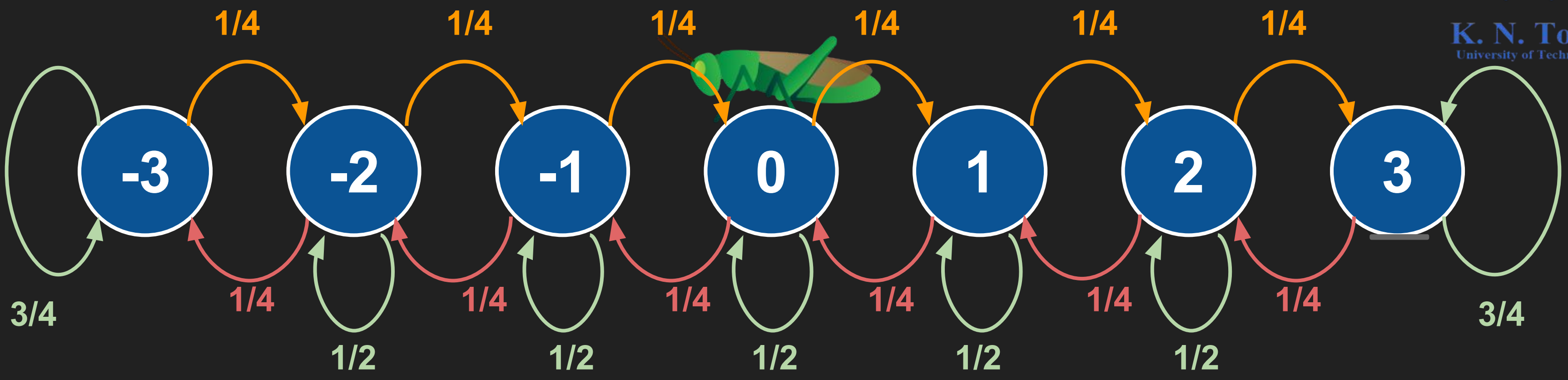
Gibbs Samplinng

Sampling Using a Markov Chain (Koller)



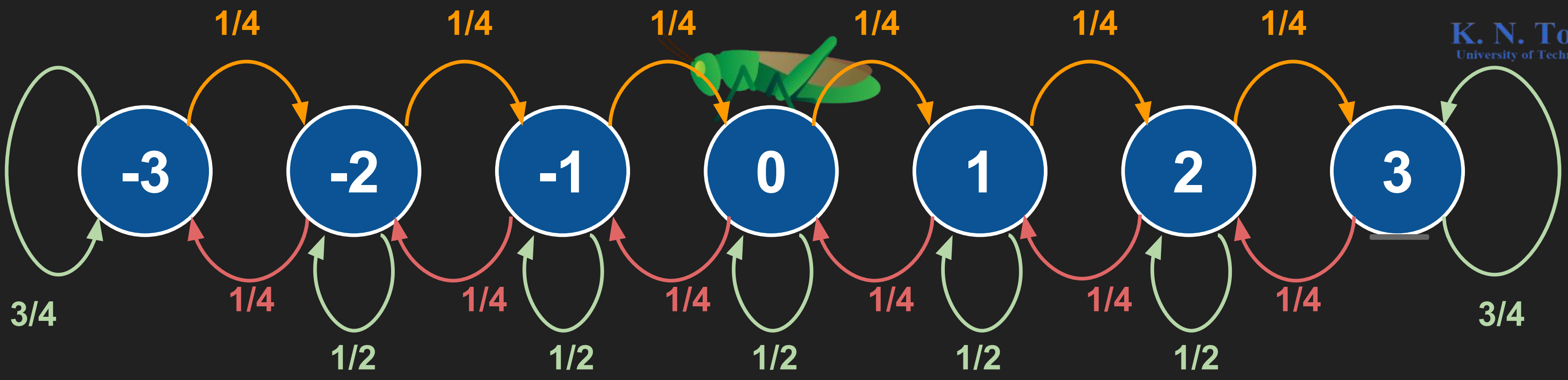
	-3	-2	-1	0	1	2	3
t=0	0	0	0	1	0	0	0
t=1							
t=2							

Sampling Using a Markov Chain (Koller)



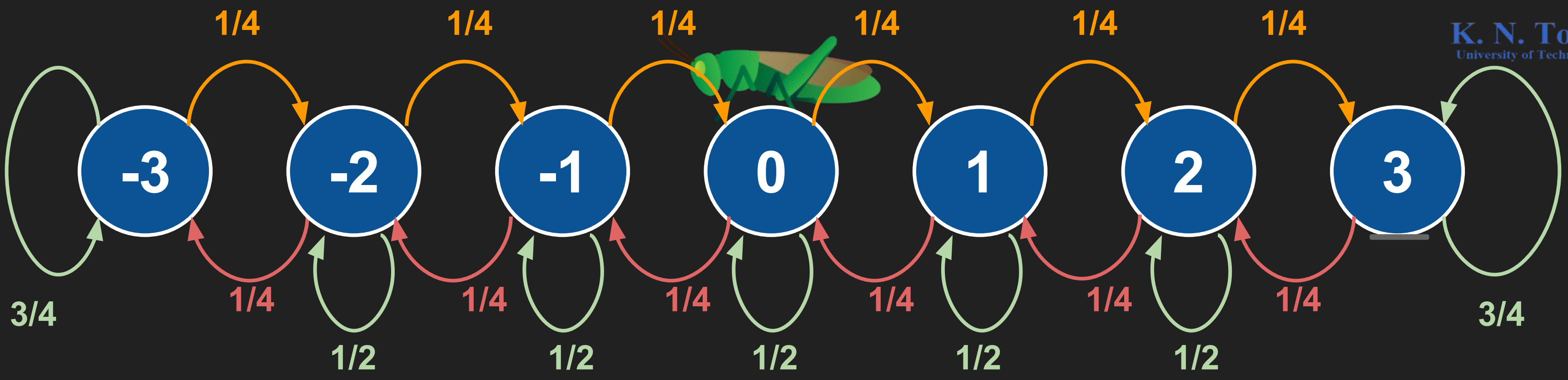
	-3	-2	-1	0	1	2	3
t=0	0	0	0	1	0	0	0
t=1	0	0	1/4	1/2	1/4	0	0
t=2							

Sampling Using a Markov Chain (Koller)



	-3	-2	-1	0	1	2	3
t=0	0	0	0	1	0	0	0
t=1	0	0	1/4	1/2	1/4	0	0
t=2	0	1/16	1/4	3/8	1/4	1/16	0

Sampling Using a Markov Chain (Koller)



	-3	-2	-1	0	1	2	3
t=0	0	0	0	1	0	0	0
:	:	:	:	:	:	:	:
$t \to \infty$	1/7	1/7	1/7	1/7	1/7	1/7	1/7

Sampling Using a Markov Chain

	-3	-2	-1	0	1	2	3
$P^{t=0}(x)$	0	0	0	1	0	0	0
$P^{t=1}(x)$	0	0	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	0	0
$P^{t=2}(x)$	0	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{16}$	0
$P^{t=\infty}(x)$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{7}$

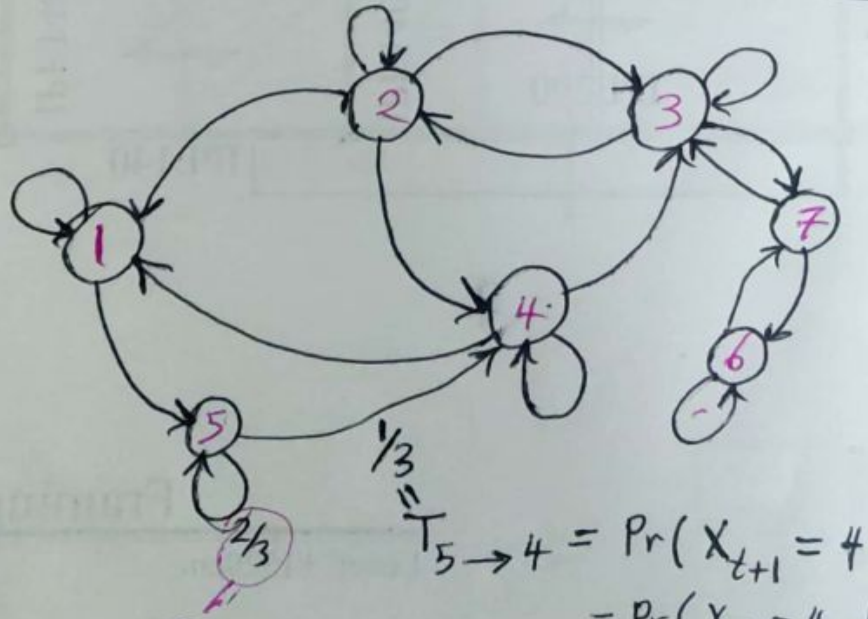
$\frac{1}{4} \times \frac{1}{4}$

$\frac{1}{4} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{4}$

$\frac{1}{2} \times \frac{1}{2} + \frac{1}{4} \times \frac{1}{4} + \frac{1}{4} \times \frac{1}{4}$

$\frac{1}{4} + \frac{1}{16} + \frac{1}{16}$

State Transition Graph



$T_{5 \to 5}$

$P(X)$

$X \in \{1,2,3,4,5,6,7\}$

X_t = ~~stat~~ state ~~curren~~ at time t

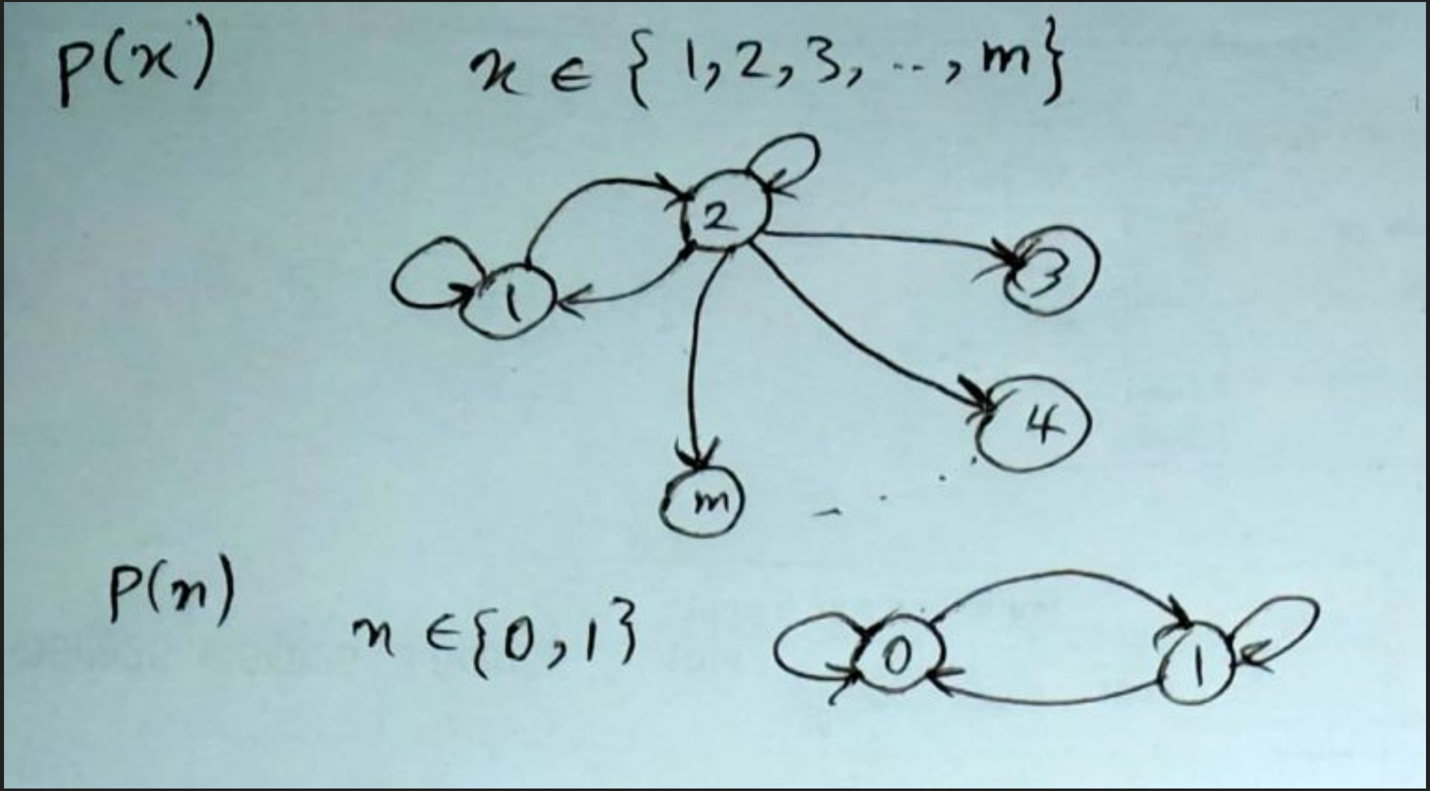
$$\frac{1}{3} = T_{5 \to 4} = Pr(X_{t+1} = 4 \mid X_t = 5)$$

$$= Pr(X_t = 4 \mid X_{t-1} = 5)$$

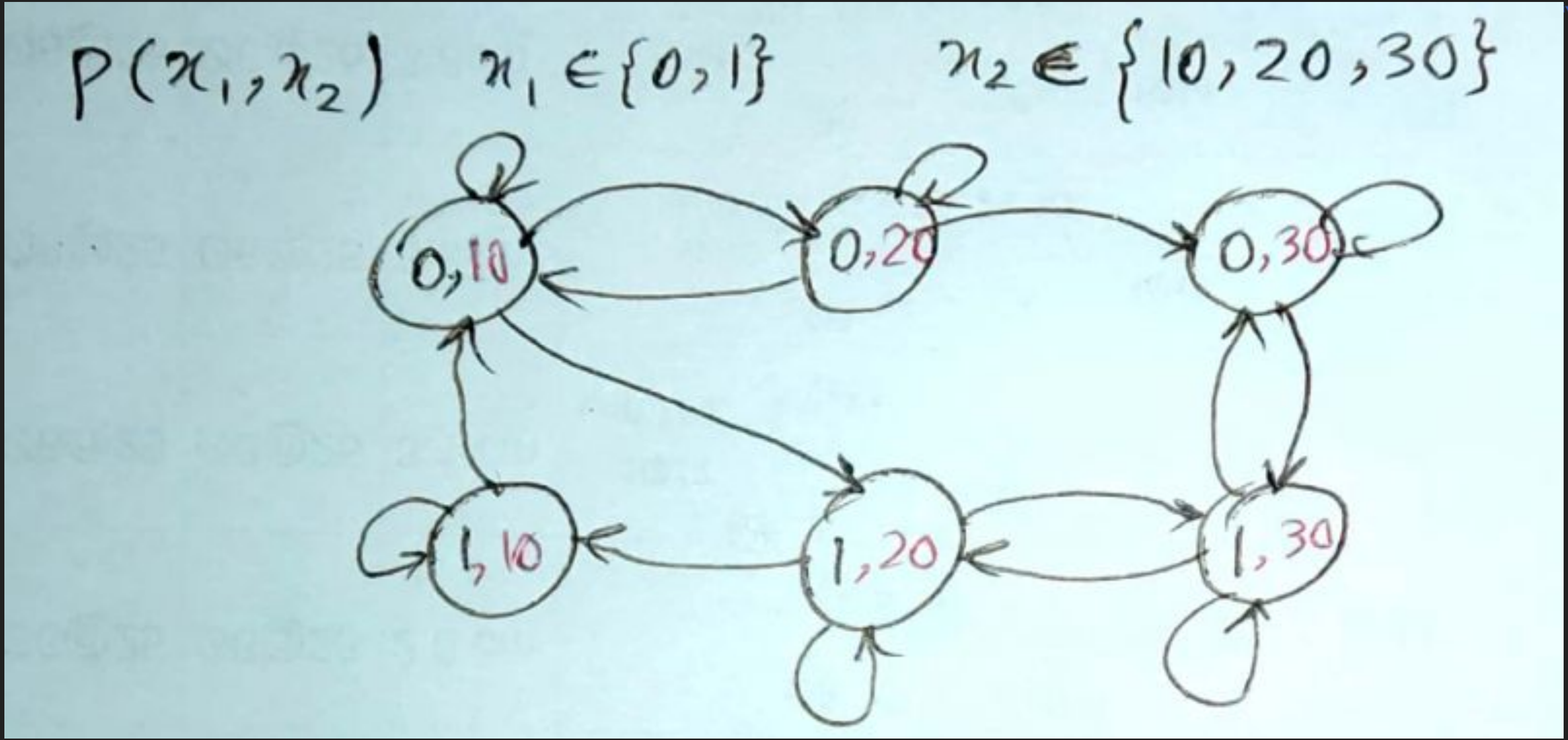
independent of t

$$Pr(X_t = a) = \sum_{b=1}^{7} Pr(X_t = a \mid X_{t-1} = b)\, p(X_t = b)$$

State Transition Graph - Single Variable



State Transition Graph - Two Variables

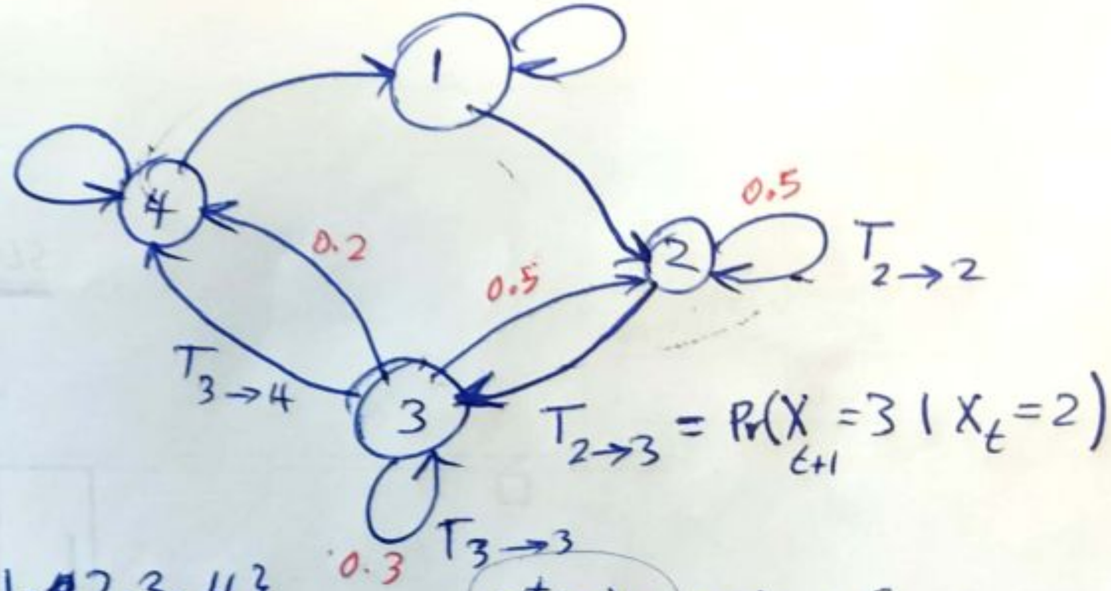


State Transition Graph - Multiple Variables

$P(x_1, x_2, \ldots, x_n\}$ & $x_i \in \{0, 1\}$

$\Longrightarrow$ there are 2^n states in the state transition graph.

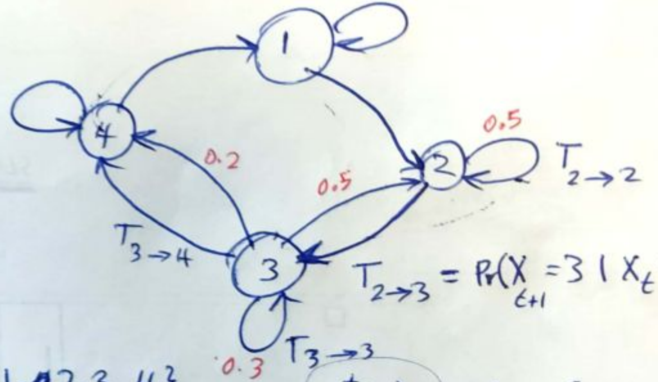
State Transition Graph



$X \in \{1, 2, 3, 4\}$

$P^t(X)$: prob. of being in state X at time t

State Transition Graph

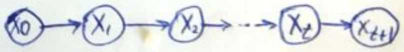


$T_{2\to3} = \Pr(X_{t+1} = 3 \mid X_t = 2)$

$X \in \{1, 2, 3, 4\}$

$p^t(X)$: prob. of being in state X at time t

$p^t(X) = \sum$



$$p^t(a) = \sum_b \Pr(X_{t+1} = a \mid X_t = b)\, p^{t-1}(b)$$

$$p^t(a) = \sum_b T_{b\to a}\, p^{t-1}(b) \qquad a = 1, 2, 3, 4$$

Stationary Distribution

$$p^t(a) = \sum_b \Pr(X_{t+1}=a \mid X_t=b)\, p^{t-1}(b)$$

$$p^t(a) = \sum_b T_{b\to a}\, p^{t-1}(b) \qquad a=1,2,3,4$$

$$\lim_{t\to\infty} p^t(a) = p^{\infty}(a) = \pi(a) = \pi_a$$

if converge:

$$\lim_{t\to\infty} p^t(a) = \lim_{t\to\infty} \sum_b T_{b\to a}\, p^{t-1}(b)$$

$$\boxed{\pi(a) = \sum_b T_{b\to a}\, \pi(b)} \qquad a=1,2,3,4$$

Solving for the Stationary Distribution

$$\pi_a = \sum_b T_{b\to a}\,\pi_b \qquad a = 1, 2, 3, 4$$

$$\pi_1 = \sum_b T_{b\to 1}\,\pi_b$$

$$\text{I}\quad \pi_1 = T_{1\to1}\,\pi_1 + T_{2\to1}\,\pi_2 + T_{3\to1}\,\pi_3 + T_{4\to1}\,\pi_4$$

$$\text{II}\quad \pi_2 = T_{1\to2}\,\pi_1 + T_{2\to2}\,\pi_2 + T_{3\to2}\,\pi_3 + T_{4\to2}\,\pi_4$$

$$\text{III}\quad \pi_3 = T_{1\to3}\,\pi_1 + T_{2\to3}\,\pi_2 + T_{3\to3}\,\pi_3 + T_{4\to3}\,\pi_4$$

$$\text{IV}\quad \pi_4 = T_{1\to4}\,\pi_1 + T_{2\to4}\,\pi_2 + T_{3\to4}\,\pi_3 + T_{4\to4}\,\pi_4$$

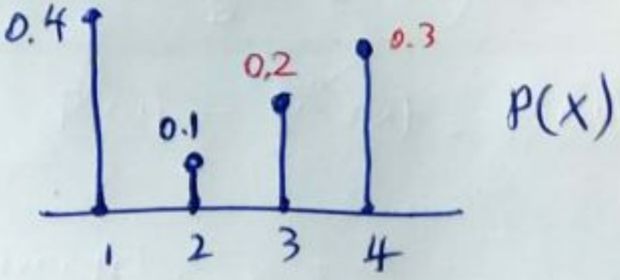
4 equation 4 unknows $(\pi_1, \pi_2, \pi_3, \pi_4)$

3 independent equations

$$\text{V}\quad \boxed{\pi_1 + \pi_2 + \pi_3 + \pi_4 = 1}$$

Solving for the Stationary Distribution

Assume that $\pi_1 = 0.4$ $\pi_2 = 0.1$ $\pi_3 = 0.2$ $\pi_4 = 0.3$



Stationary Distribution

State Transition Graph - Two Variables

$$p_j^t = P(S_t = j) = \sum_{k=1}^{m} P(S_t = j, S_{t-1} = k)$$

$$= \sum_{k=1}^{m} P(S_t = j \mid S_{t-1} = k)\, P(S_{t-1} = k)$$

$$p_j^t = \sum_{k=1}^{m} T_{k \to j} \; p_k^{t-1}$$

$$p^t = (p_1^t, p_2^t, \ldots, p_m^t)$$

$$\lim_{t \to \infty} p^t = ?$$

exists? NOT Always

what is it?

if it converges

Stationary Distribution

$$\lim_{t \to \infty} \quad P_j^t = \sum_{k=1}^{m} T_{k \to j} \, P_k^{t-1}$$

$$P_j^\infty = \sum_{k=1}^{m} T_{k \to j} \, P_k^\infty$$

$$\boxed{\pi_j = \sum_{k=1}^{m} T_{k \to j} \, \pi_k \qquad j = 1, 2, \ldots, m}$$

Stationary Distribution

$$\boxed{\pi_j = \sum_{k=1}^{m} T_{k\to j} \cdot \pi_k \qquad j = 1, 2, \ldots, m}$$

Assume that $T_{k\to j}$ is know for all $k, j \in \{1, \ldots, m\}$

Gives $\underline{m}$ equations in $\underline{m}$ unknows $\pi_1, \ldots, \pi_m$

there are only $\underline{m-1}$ <u>independent</u> equations.

$$1 = \sum_{j=1}^{m} \pi_j = \sum_{j=1}^{m}\sum_{k=1}^{m} T_{k\to j} \cdot \pi_k = \underbrace{\sum_{k=1}^{m} \underbrace{\sum_{j=1}^{m} T_{k\to j}}_{1} \pi_k}_{1} = 1$$

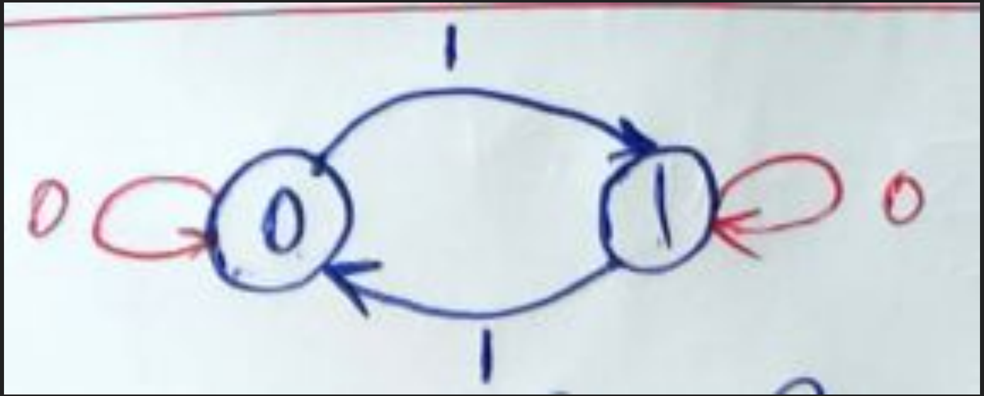
$$\left.\begin{cases} m-1 \text{ independent equation} \\ \quad + \\ \pi_1 + \pi_2 + \cdots + \pi_n = 1 \end{cases}\right\} \Rightarrow \text{solve for } \pi_1, \pi_2, \ldots, \pi_n$$

Stationary Distribution

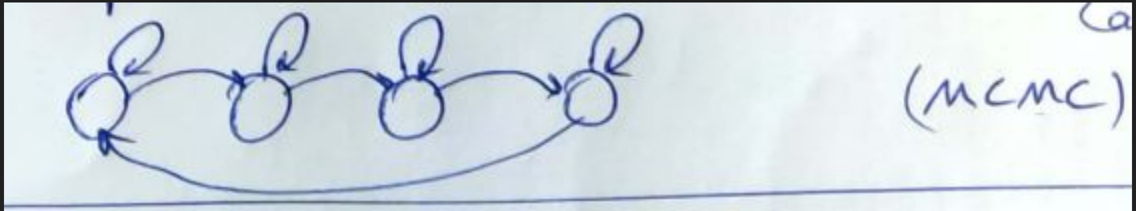
$\Longrightarrow$ we have a distribution $P_j^{\infty} = P(X_{t=\infty} = j)$

↪ if we move long enough in the state transition graph we have a sample from $P(X_{t=\infty} = j) = P^{\infty}$

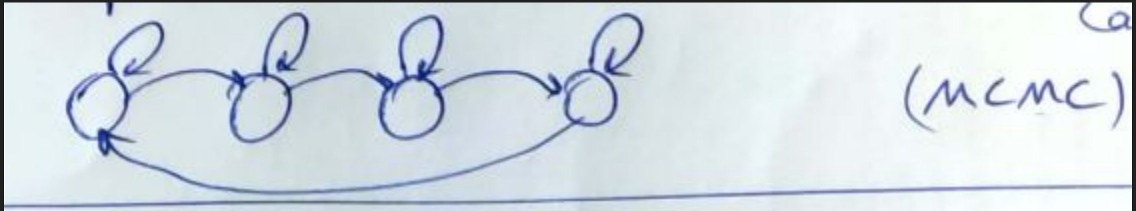
Convergence is not always guaranteed



MCMC Convergence



MCMC Convergence



Stationary Distribution

$\Longrightarrow$ we have a distribution $P_j^{\infty} = P(X_{t=\infty} = j)$

$\hookrightarrow$ if we move long enough in the state transition graph we have a sample from $P(X_{t=\infty} = j) = P^{\infty}$

Stationary Distribution

Main idea in MCMC:

→ we have a distribution $P(x)$

→ create a state transition graph (explicit/implicit)

a set $T_{i \to j}$ such that $P(x=k) = P^{\infty}(X=k)$

(now we have π_k's in $\pi_j = \sum_{j=1}^{m} T_{k \to j}\, \pi_k$ &

we need to set $T_{k \to j}$)

More in the state transition graph long enough to have a sample from $p(x)$.

Gibbs Sampling

$$a^{t+1} \sim P(A \mid B=b^t, C=c^t, D=d^t, E=e^t)$$

$$b^{t+1} \sim P(B \mid A=a^{t+1}, C=c^t, D=d^t, E=e^t)$$

$$c^{t+1} \sim P(C \mid A=a^{t+1}, B=b^{t+1}, D=d^t, E=e^t)$$

$$d^{t+1} \sim P(D \mid A=a^{t+1}, B=b^{t+1}, C=c^{t+1}, E=e^t)$$

$$e^{t+1} \sim P(E \mid A=a^{t+1}, B=b^{t+1}, C=c^{t+1}, D=d^{t+1})$$

$$(a^t, b^t, c^t, d^t, e^t)$$

$$(a^{t+1}, b^{t+1}, c^{t+1}, d^{t+1}, e^{t+1})$$

Computing conditionals

$$P(A \mid B, C, D, E) = \frac{P(A, B, C, D, E)}{\sum_{A} P(A, B, C, D, E)}$$

Computing conditionals

K. N. Toosi University of Technology

$$P(A,B,C,D,E) = \frac{1}{Z}\,\phi_1(A)\,\phi_2(A,B)\,\phi_3(B,C)\,\phi_4(C,D,E)$$

$$P(A \mid b,c,d,e) = \frac{\cancel{\frac{1}{Z}}\,\phi_1(A)\,\phi_2(A,b)\,\phi_3(b,c)\,\phi_4(c,d,e)}{\sum_{a} \cancel{\frac{1}{Z}}\,\phi_1(a)\,\phi_2(a,b)\,\phi_3(b,c)\,\phi_4(c,d,e)}$$

$$P(A \mid b,c,d,e) = \frac{\cancel{\frac{1}{Z}}\,\phi_1(A)\,\phi_2(A,b)\,\cancel{\phi_3(b,c)}\,\cancel{\phi_4(c,d,e)}}{\sum_{d} \cancel{\frac{1}{Z}}\,\phi_1(a)\,\phi_2(a,b)\,\cancel{\phi_3(b,c)}\,\cancel{\phi_4(c,d,e)}}$$

$$= \frac{\phi_1(A)\,\phi_2(A,b)}{\sum_{a} \phi_1(a)\,\phi_2(a,b)}$$

Warm-up (burn-in) phase and mixing