

# Chapman Kolmogorov Equations for DTMC

- ▶ Consider a Markov coin and its transition probability matrix

$$P = \begin{bmatrix} p_{1,1} & p_{1,-1} \\ p_{-1,1} & p_{-1,-1} \end{bmatrix}.$$

- ▶ Given  $X_0 = 1$ , what is  $P(X_2 = 1)$ ?

$$\begin{aligned} P(X_2 = 1|X_0 = 1) &= P(X_2 = 1|X_1 = 1, X_0 = 1)P(X_1 = 1|X_0 = 1) \\ &+ P(X_2 = 1|X_1 = -1, X_0 = 1)P(X_1 = -1|X_0 = 1) \\ &= p_{1,1}^2 + p_{-1,1}p_{1,-1} \end{aligned}$$

- ▶ Here the first inequality follow from the fact that

$$P(C|A) = P(C|BA)P(B|A) + P(C|B^cA)P(B^c|A) \text{ HW: Verify}$$

- ▶ Similarly,  $P(X_2 = -1|X_0 = 1)$ ,  $P(X_2 = 1|X_0 = -1)$ ,  $P(X_2 = -1|X_0 = -1)$  can be obtained and these are elements of a two-step transition matrix  $P^{(2)}$ .

# Chapman Kolmogorov Equations for DTMC

- ▶ The two step transition probability matrix  $P^{(2)}$  is given by
$$P^{(2)} = \begin{bmatrix} p_{1,1}^2 + p_{1,-1}p_{-1,1} & p_{1,1}p_{1,-1} + p_{1,-1}p_{-1,-1} \\ p_{-1,1}p_{1,1} + p_{-1,-1}p_{-1,1} & p_{-1,1}p_{1,-1} + p_{-1,-1}^2 \end{bmatrix}.$$
- ▶ This implies that  $P^{(2)} = P \times P = P^2$ .
- ▶ In general,  $P^{(n)} = P^n$ .
- ▶ Chapman-Kolmogorov equations are a further generalization of this.

$$P^{(n+l)} = P^{(n)} P^{(l)}$$

- ▶ We wont see the proof of this.

# RECAP

- ▶ Markov property:

$$P(X_n = j | X_1 = x_1, \dots, X_{n-1} = x_{n-1}) = P(X_n = j | X_{n-1} = x_{n-1})$$

- ▶ You need  $\bar{\mu} = (\mu_1, \dots, \mu_M)$  and t.p.m  $P$  to write down the finite dimensional distributions

$$P(X_0 = x_0, X_1 = x_1, \dots, X_k = x_k) = p_{x_{k-1}, x_k} \times \dots \times p_{x_0, x_1} \mu_{x_0}$$

- ▶ Chapman Kolmogorov Equations tell us that the n-step t.p.m is just  $P^n$ .

# Classification of states

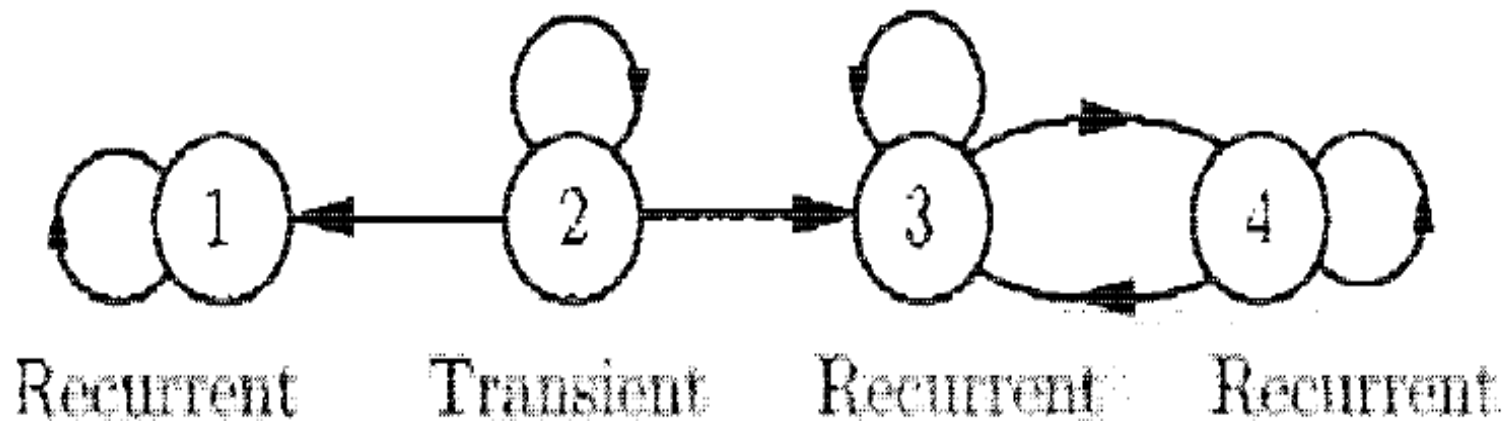
- ▶ Consider a Markov process with state space  $\mathcal{S}$
- ▶ We say that  $j$  is accessible from  $i$  if  $p_{ij}^n > 0$  for some  $n$ .
- ▶ This is denoted by  $i \rightarrow j$ .
- ▶ if  $i \rightarrow j$  and  $j \rightarrow i$  then we say that  $i$  and  $j$  communicate. This is denoted by  $i \leftrightarrow j$ .

A chain is said to be irreducible if  $i \leftrightarrow j$  for all  $i, j \in \mathcal{S}$ .

- ▶ Are the examples of Markovian coin and dice we have considered till now irreducible? **check!**

# Recurrent and Transient states

- ▶ We say that a state  $i$  is recurrent if  $F_{ii} = P(\text{ ever returning to } i \text{ having started in } i ) = 1$ .
- ▶  $F_{ii}$  is not easy to calculate. (Not part of this course)
- ▶ If a state is not recurrent, it is transient.
- ▶ For a transient state  $i$ ,  $F_{ii} < 1$ .
- ▶ If  $i \leftrightarrow j$  and  $i$  is recurrent, then  $j$  is recurrent.



# Limiting probabilities

$$\blacktriangleright P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0.6 & 0.4 \\ 0.6 & 0.4 & 0 \end{bmatrix} \quad P^5 = \begin{bmatrix} .06 & .3 & .64 \\ .18 & .38 & .44 \\ .38 & .44 & .18 \end{bmatrix} \quad P^{30} = \begin{bmatrix} .23 & .385 & .385 \\ .23 & .385 & .385 \\ .23 & .385 & .385 \end{bmatrix}$$

$$\blacktriangleright P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} \quad \lim_{n \rightarrow \infty} P^n = \begin{bmatrix} \frac{b}{a+b} & \frac{a}{a+b} \\ \frac{b}{a+b} & \frac{a}{a+b} \end{bmatrix}$$

- ▶ What is the interpretation of  $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = [\lim_{n \rightarrow \infty} P^n]_{ij}$ ?
- ▶  $\alpha_j = \lim_{n \rightarrow \infty} p_{ij}^{(n)}$  denotes the probability of being in state  $j$  after a large time from starting in state  $i$ .
- ▶ For an  $M$  state DTMC,  $\alpha = (\alpha_1, \dots, \alpha_M)$  denotes the limiting distribution.
- ▶ How do we obtain the limiting distribution  $\alpha$ ? Does it always exist?

# Stationary distribution

The **stationary distribution** of a Markov chain is defined as a solution to the equation  $\pi = \pi P$ .

- ▶  $\pi P$  is essentially the p.m.f of  $X_1$  having picked  $X_0$  according to  $\pi$ .
- ▶  $\pi = \pi P$  says that, if the initial distribution is  $\pi$ , then the distribution of  $X_1$  is also  $\pi$ .
- ▶ Continuing this argument, the p.m.f of  $X_n$  for any  $n$  is  $\pi$  and there is no dependence on the starting state.
- ▶ MCMC algorithms use this idea (at stationarity successive states of the Markov chain have p.m.f  $\pi$ ) to sample from target distribution  $\pi$ .

# Limiting vs Stationary distribution

- ▶ Obtain stationary distribution for the Markov Chain with transition probability  $P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0.6 & 0.4 \\ 0.6 & 0.4 & 0 \end{bmatrix}$
- ▶ The limiting distribution  $\alpha$  need not exist for some Markov chains, but the stationary distribution  $\pi$  exists. For example for  $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ .
- ▶ The limiting distribution if it exists, is same as the stationary distribution, i.e.  $\alpha_i = \pi_i$  for all  $i$ .