(1)

LECTURE 4: INTRODUCTION TO MARKOV CHAINS

Def: Let $\{X_n; n=0,1,..\}$ be a stochastic process on $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{F}_n = \sigma(X_0, .. X_n)$ be the natural filtration. $\{X_n\}$ is called a Markov chain if for any borel set $B \in \mathcal{B}(\mathbb{R})$:

$$\mathbb{P}(X_{n+1} \in B \mid \mathcal{F}_n) = \mathbb{P}(X_{n+1} \in B \mid X_n) = \mathbb{P}(X_{n+1} \in B \mid \sigma(X_n))$$

(recall that a conditional probability is a random variable and it only depends on the conditioning σ-algebra).

Physical Interpretation: the future evolution of the process depends only on the current state of the system and it is independent of the past.

Def: A Markov chain (MC) $\{X_n\}$ that takes values on a discrete space S is called a "discrete" MC. That is, $\forall n \in \mathbb{N}$, $X_n \in S$ (wlog $S = \mathbb{Z}$).

For a discrete Markov chain, the conditional probabilities can be described as matrices (finite when S is finite):

$$P_{ij}(n) \equiv \mathbb{P}\{X_{n+1} = j \mid X_n = i\}$$

For a general MC,

$$P_i(n; dx) = \mathbb{P}(X_{n+1} \in dx \mid X_n = i)$$

defines a density, called the transition density, or transition kernel.

Def: Let $\{X_n\}$ be a Markov chain on $(\Omega, \mathcal{F}, \mathbb{P})$. If the transition kernel is independent of n, that is:

$$\mathbb{P}(X_{n+1} \in B \mid \mathcal{F}_n) = \mathbb{P}(X_{m+1} \in B \mid \mathcal{F}_m)$$

$\forall n, m \in \mathbb{N}$, then we say that $\{X_n\}$ is a (time) homogeneous Markov chain. Otherwise it is called non-homogeneous.

Example: Ross p.193 (students to read all examples in texts)

Probability model $\Omega = \{$☀, 🌧$\}$ (rain or no rain)

Historical data fits the following model:

Yesterday	Today	$\mathbb{P}$(Rain tomorrow)
🌧	🌧	0.7
☀	🌧	0.5
🌧	☀	0.4
☀	☀	0.2

Let $X_n = \begin{cases} 0 & \text{if rain on day } n \\ 1 & \text{if no rain on day } n \end{cases}$

Is $\{X_n\}$ a Markov chain?

"Markovianising" a process: enlarge the state to contain enough information into the past.

Define: $Y_n = \begin{cases} 0 & \text{if } (0,0) \\ 1 & \text{if } (0,1) \\ 2 & \text{if } (1,0) \\ 3 & \text{if } (1,1) \end{cases}$ Is $\{Y_n\}$ a MC? Is it homogeneous?

Consider a finite homogeneous MC $\{X_n\}$. Then the process is completely defined once the transition probability p_{ij} and the initial distribution of X_0 are specified. This means that given $\{p_{ij},\ i,j\in S\}$ and $\mathbb{P}(X_0=i)\ \forall i\in S$, the distribution of $X_n\ \forall n\geq 0$ is well defined. (p.97 TK)

Notation: For a homogeneous MC,

$$P_{ij}^{(m)} = \mathbb{P}\{X_{n+m}=j \mid X_n=i\}$$ is called

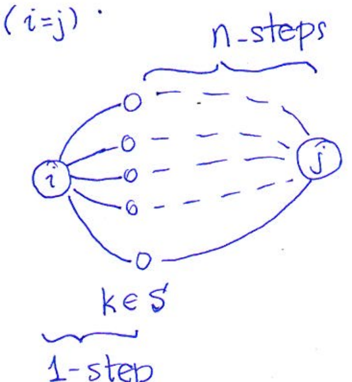
the m-step transition probability.

Theorem: Chapman-Kolmogorov equations. The n-step transition probabilities of a homogeneous Markov chain satisfy:

$$P_{ij}^{(n)} = \sum_{k=0}^{\infty} P_{ik}P_{kj}^{(n-1)}$$

where $P_{ij}^{(0)} = \mathbb{1}(i=j)$.

Visualization:



Models (TK p. 105 - 112)
a) Inventory
b) Ehrenfest urn model
c) Genetics
d) Queuing

Proof:

$$\mathbb{P}(X_{n+1}=j \mid X_1=i) = \sum_{k\in S}\mathbb{P}(X_{n+1}=j \mid X_2=k, X_1=i)\times \mathbb{P}(X_2=k\mid X_1=i)$$

(emphasis on conditioning)

~~Example~~

Stopped Markov Chains (Ross p. 200)

$\{X_n\}$ a MC with transition probabilities P_{ij}.

Let $\mathcal{A}$ be a "target" set of states $\mathcal{A}\subset S$.

$$N = \min(n : X_n \in \mathcal{A})$$

is a random stopping time w.r.t. the natural filtration $\{\sigma(X_0,\dots X_n)\}$.

We want to calculate

$$\beta = \mathbb{P}(N<\infty)$$

probability that the target set is attained by the process.

Example: probability of winning the game of craps

$$W_n = \begin{cases} X_n, & n<N \\ A & n\geq N\end{cases}$$

The MC $\{W_n\}$ is called a "stopped" MC, with state space $\{S\setminus\mathcal{A}\}\cup\{A\}$. Transition probabilities are

$$Q_{ij} = P_{ij} \quad i\notin\mathcal{A},\ j\notin\mathcal{A}$$

$$Q_{i,A} = \sum_{j\in\mathcal{A}} P_{ij} \quad i\notin\mathcal{A}$$

$$Q_{A,A} = 1$$

Using the stopped chain,

$$\mathbb{P}(N\leq m) = \mathbb{P}(W_m = A)$$

and $\mathbb{P}(W_m=A\mid X_0=i) = Q_{i,A}^{(m)}$. See applications in Ross examples 4.12, 4.13 and end of section.

CLASSIFICATION OF STATES

Def: Let $\{X_n\}$ be a MC on $(\Omega, \mathcal{F}, \mathbb{P})$ with state space $S \subset \mathbb{N}$. State j is said to be accessible from i if $P_{ij}^{n} > 0$ for some integer $n \geq 0$. Notation: $i \to j$

P.204 (Ross), P.234 TK.

Def: Two states $i, j \in S$ are said to communicate if $i \to j$ and $j \to i$. Notation: $i \leftrightarrow j$.

If two states i and $j \in S$ do not communicate then it follows that either $P_{ij}^{(n)} = 0 \ \forall n \geq 0$ or $P_{ji}^{(n)} = 0 \ \forall n \geq 0$.

Theorem: Communication is an equivalence relationship.

Proof:

(i). $P_{ii}^{(0)} = 1$ by definition

(ii). $i \leftrightarrow j \iff j \leftrightarrow i$ by definition

(iii). If $i \leftrightarrow j$ and $j \leftrightarrow k$ then $\exists\, n, m$ such that $P_{ij}^{(n)} > 0$ and $P_{jk}^{(m)} > 0$. Using Chapman Kolmogorov equations:

$$P_{ik}^{n+m} = \sum_{q \in S} P_{iq}^{(n)} P_{qk}^{(m)} = \underbrace{\sum_{q \neq j} P_{iq}^{(n)} P_{qk}^{(m)}}_{\geq 0} + \underbrace{P_{ij}^{(n)} P_{jk}^{(m)}}_{>0}$$

which shows that $i \to k$. A symmetric argument show that $k \to i$ as well, proving the result.

Examples: games played in class.

GAME 1 (build with students):

$$\begin{pmatrix} 2/5 & 0 & 2/5 & 1/5 \\ 0 & 1 & 0 & 0 \\ 3/8 & 0 & 1/2 & 1/8 \\ 1/2 & 1/8 & 3/8 & 0 \end{pmatrix}$$

GAME 2

$$\begin{pmatrix} 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1/3 & 2/3 \\ 2/5 & 3/5 & 0 & 0 \\ 1/8 & 0 & 7/8 & 0 \end{pmatrix}$$

Diagrams and structure: notice that communication and ensuing classes are independent of the numerical values of P_{ij}, only whether $P_{ij} > 0$ or $P_{ij} = 0$. ☒☒☒

Let $\tau = \min(n : X_n = X_0)$ be the first return time to the initial state (a stopping time).

Def: $f_i^{(n)}$ is ~~called~~ the probability that the first return to state i happens at the n-th transition:

$$f_i^{(n)} = \mathbb{P}(\tau = n \mid X_0 = i)$$

[P.239-241 TK and 205-206 Ross]

Notice that $f_i^{(1)} = P_{ii}^{(1)}$

$$P_{ii}^{(n)} = \sum_{k=0}^{n} f_i^{(k)} P_{ii}^{(n-k)}, \quad n \geq 1. \quad \text{(TK 240 for proof)}$$

Def: Let $f_i = \mathbb{P}(\tau < \infty \mid X_0 = i)$ denote the probability that, starting at i, the process eventually returns to i:

$$f_i = \sum_{n\geq 1} \mathbb{P}(\tau = n \mid X_0 = i) = \sum_{n \geq 1} f_i^{(n)} = \lim_{N\to\infty} \sum_{n=1}^{N} f_i^{(n)}.$$

Def: A state $i \in \mathcal{S}$ is called:

- recurrent if $f_i = 1$,
- transient if $f_i < 1$.

Let i be a transient state so that $f_i < 1$. In this case, given a return to i, say $X_n = i$, the probability that it returns to i again is f_i because of the Markov property. Therefore,

$$\mathbb{P}(1 \text{ visit}) = f_i(1-f_i)$$

$$\mathbb{P}(2 \text{ visits}) = f_i^2(1-f_i)$$

$$\vdots$$

$$\mathbb{P}(k \text{ visits}) = f_i^k(1-f_i)$$

$$\begin{tikzcd} i \arrow[r, bend left, "f_i"] & i \arrow[r, bend left, "f_i"] & i \end{tikzcd}$$

"success" when return to i

Let M count the total number of visits (return) to state i, that is:

$$M = \sum_{n=1}^{\infty} \mathbb{1}(X_n = i)$$

The random variable M satisfies $\mathbb{P}(M=k) = f_i^k(1-f_i) \Rightarrow$ has geometric distribution with parameter $f_i < 1$.

Theorem: A state i is recurrent iff $\sum_{n\geq 1} P_{ii}^{(n)} = +\infty$ and transient if $\sum_{n=1}^{\infty} P_{ii}^{(n)} < \infty$. p. 206 Ross, p. 241 TK

Proof: If i is transient then $f_i < 0$. We know that in this case $\mathbb{E}(M \mid X_0 = i) = \frac{f_i}{1-f_i} < \infty$ is finite. From the definition of M it follows that

$$\mathbb{E}(M \mid X_0 = i) = \mathbb{E}\Big(\sum_{n\geq 1} \mathbb{1}(X_n = i) \,\Big|\, X_0 = i\Big)$$

$$= \sum_{n\geq 1} P_{ii}^{(n)} < \infty.$$

On the other hand if i is recurrent then $f_i = 1$ and the process will visit i infinitely often: $\mathbb{P}(X_m = i, \text{ some } m \geq n \mid X_n = i) = 1$. In this case $\mathbb{E}M = +\infty \Leftrightarrow$

$$\mathbb{E}\Big(\sum_{n\geq 1} \mathbb{1}(X_n = i) \,\Big|\, X_0 = i\Big) = +\infty \Leftrightarrow$$

$$\sum_{n\geq 1} P_{ii}^{(n)} = +\infty.$$

QED

Def: A recurrent state i such that $P_{ii} = 1$ is called an absorbing state.

[Go back to examples in games]. Ross p. 208 random walk, TK p. 241 random walk

Markov chain analysis $\nearrow$ absorption probabilities $\searrow$ recurrent classes: steady-state

Theorem: Recurrence is a class property.

Def: The period $d(i)$ of state i is the greatest common divisor of all integers $n \geq 1$ for which $P_{ii}^{(n)} > 0$. If $P_{ii}^{(n)} = 0 \ \forall n \Rightarrow d(i) = \infty$

Theorem: Periodicity is a class property [TK p. 239].

Def: A recurrent state $i \in \mathcal{S}$ is called positive recurrent if the expected return time is finite: $\mathbb{E}(\tau \mid X_0 = i) < \infty$. [Random walk Ross]