REVIEW OF MARKOV CHAINS (Stationary Analysis)

(A)

Def: Let $\{X_n ; n=0,1,2,..\}$ be a stochastic process on $(\Omega,\mathcal{F},\mathbb{P})$ and let $\mathcal{F}_n = \sigma(X_0,..,X_n)$ be the natural filtration of the process. $\{X_n\}$ 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 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 current state and it is independent of the past: memoryless.

Notation: the state space S is usually $S\subset\mathbb{R}$, but it can also be in several dimensions. To write the density of the state X_{n+1} given the current state X_n we can use the notation:

$$\mathbb{P}(X_{n+1}\in dx\mid\mathcal{F}_n) = P(X_n, dx).$$

and $P(x,dx)$ is called the transition kernel.

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

For a discrete MC the kernels can be described by matrices:

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

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

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

then we say that $\{X_n\}$ is a (time) homogeneous Markov chain. Otherwise it is called non-homogeneous.

Classification of States

In this section we consider discrete, homogeneous MCs. Most examples use a finite state space S.

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 $P^n_{ij}>0$ for some integer n.

Notation: $i\to j$.

[$P^n_{ij} = \mathbb{P}(X_{\ell+n}=j\mid X_\ell=i)$ is the n-step transition probability].

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 $i,j\in S$ do not communicate then it follows that either $P^n_{ij}=0$ $\forall n\geq 0$ or $P^n_{ji}=0$ $\forall n\geq 0$.

Theorem: Communication is an equivalence relationship. $\boxed{\text{CLASSES}}$

Proof:

(i). $P^0_{ii}=1$ by definition (reflexive)

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

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

$$P^{n+m}_{ik} = \sum_{q\in S} P^n_{iq}P^m_{qk} = \sum_{q\neq j} P^n_{iq}P^m_{qk} + \underbrace{P^n_{ij}P^m_{jk}}_{>0} > 0 \Rightarrow i\leftrightarrow k.$$

B

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

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

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

Result: $f_i^{(1)} = P_{ii}$, and $P_{ii}^n = \sum_{k=0}^{n} f_i^{(k)} P_{ii}^{n-k}$, $n \geq 1$.

Def. $f_i = \mathbb{P}(\tau < \infty \mid X_0 = i)$ is the probability that the process eventually returns to the initial state i. It satisfies:

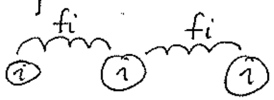
$$f_i = \sum_{n\geq 1} \mathbb{P}(\tau = n \mid X_0 = i) = \sum_{n \geq 1} f_i^{(n)} = \lim_{N\to\infty} \sum_{n=1}^{N} f_i^{(n)}.$$

Def: A state $i \in S$ is called:

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

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

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



"succes" of Bernoulli trial when return to i

Let M count the number of visits to state i, then:

$$M = \sum_{n=1}^{\infty} \mathbb{1}(X_n = i) \text{ is a random variable } \sim \text{Geom}(f_i).$$

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

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

Markov chain analysis
- → absorption probabilities
- → stationary analysis (recurrent classes)

Theorem: Recurrence is a class property.

Def: The period of state i is the gcd of all integers $n \geq 1$ for which $P_{ii}^n > 0$. If $P_{ii}^n = 0 \; \forall n \Rightarrow i$ has period 0.

Theorem: Periodicity is a class property.

Def. A recurrent state $i \in S$ is called positive recurrent if $\mathbb{E}[\tau \mid X_0 = i] < \infty$.

Remark: conditioning on each possible event $\{X_n = k, k \in S\}$ we have:

$$\boxed{\mathbb{P}(X_{n+1} = j) = \sum_{i \in S} p_{ij}\, \mathbb{P}(X_n = i)} \quad (*)$$

Def: An aperiodic, recurrent state (class) is called ergodic.

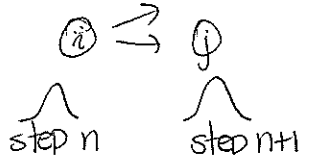
Def: A stationary measure μ of a Markov chain satisfies:

$$\mu(A) = \int_{x \in S} P(x, A)\, \mu(dx), \quad \forall A \in \mathcal{B}(\mathbb{R}).$$

If $S \subset \mathbb{N} \Rightarrow$

$$\boxed{\mu(j) = \sum_{i \in S} p_{ij}\, \mu(i), \quad \forall j \in S} \; \boxed{\sum_{i \in S} \mu(i) = 1}$$

called the "balance equation".



same distribution after one transition, but process is not constant.

c

In many examples of applications, cost functions of interest involve a long term running average. Example: calculating electricity costs per month, or wages paid to casual employees. This takes the form:

$$\lim_{N\to\infty} \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N} g(X_n)\right]$$

for some "instantaneous" cost $g: S \to \mathbb{R}$. Therefore one must understand the "limitting" behaviour of Markov chains.

Result: Suppose that for any state $i \in S$, the following limit exists:

$$\lim_{n\to\infty} \mathbb{P}(X_n = i) = \mu_i$$

Then using (*), it follows that $\{\mu_i, i\in S\}$ satisfy:

$$\mu_j = \sum_{i\in S} p_{ij}\,\mu_i,$$

that is, $\{\mu_i\}$ is a stationary measure for the Markov chain.

Remark: generally the converse is not true, that is a stationary measure may exist but not the limiting probabilities

Example. The simplest example is an alternating process (binary), where $p_{01} = p_{10} = 1$.

If $X_0 = 0$, then $X_{2k} = 0$ $\forall k \in \mathbb{N}$ and $X_{2k+1} = 1$ $\forall k \in \mathbb{N}$, so that $\mathbb{P}(X_n = 0)$ alternates between 1 and 0 $\Rightarrow$ there is no limit. However the measure $\mu(0) = \frac{1}{2} = \mu(1)$ satisfies the balance equation. [Discuss interpretation].

Proposition. Let $\{X_n\}$ be a Markov chain on a countable state space S and suppose that it has only one class (called "irreducible MC"). If the chain is ergodic then there is a unique stationary measure $\mu(\cdot)$ and

$$\lim_{n\to\infty} \mathbb{P}(X_n = i) = \mu(i), \quad \forall i \in S.$$

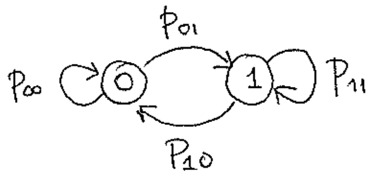
In particular, for ergodic MC and any cost function g such that $|\mathbb{E}_\mu(g)| < \infty$ (expectation w.r.t. μ) it follows that

$$\lim_{n\to\infty} \mathbb{E}[g(X_n)] = \sum_{i\in S} g(i)\mu(i)$$

$$= \lim_{N\to\infty} \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N} g(X_n)\right].$$

Example (binary MC)

Consider the following models:



The balance equations are:

$$\left.\begin{aligned} \mu_0 &= \mu_0 p_{00} + \mu_1 p_{10} \\ \mu_1 &= \mu_0 p_{01} + \mu_1 p_{11} \end{aligned}\right\} \Rightarrow \mu_0 = \frac{\mu_1 p_{10}}{(1-p_{00})} = \mu_1 \frac{p_{10}}{p_{01}} \Rightarrow$$

$$\mu_0 = (1-\mu_0)\frac{p_{10}}{p_{01}} \Rightarrow$$

$$\mu_0\left(1 + \frac{p_{10}}{p_{01}}\right) = \frac{p_{10}}{p_{01}} \quad \text{(solve linear system)}$$

This provides the stationary measure (μ_0, μ_1) as functions of the transition probabilities and can be easily evaluated.