Markov Chains

November 1, 2016

Contents

1 Markov Chain

The notes taken during the first lecture was unfortunately lost.

Theorem. (Extended Markov Property) Let X be a Markov chain, and $n \geq 0$. Let H be an event defined in terms of $X_0, ..., X_{n-1}$ (the history), and let F be an event defined in terms of $X_{n+1}, X_{n+2}, ...$ (the future). Then

$$\mathbb{P}\left(F|H, X_n = i\right) = \mathbb{P}\left(F|X_n = i\right)$$

2 Transition Probabilities

We've known that $p_{i,j}$ is the one-step transition probability,

$$\mathbb{P}\left(X_{n+1}=j|X_n=i\right)=\mathbb{P}\left(X_1=j|X_0=i\right)$$

Now it's natural to discuss the n-step transition probabilities

$$\mathbb{P}\left(X_n=j|X_0=i\right)=p_{i,j}\left(n\right)$$

Theorem. (Chapman-Kolmogorov Equations)

$$p_{i,j}\left(m+n\right)=\sum_{k\in S}p_{i,k}\left(m\right)p_{k,j}\left(n\right)$$

Proof.

$$\begin{aligned}p_{i,j}\left(m+n\right)&=\sum_{k\in S}\mathbb{P}\left(X_{m+n}=j|X_m=k\right)\mathbb{P}\left(X_m=k\right)\\&=\sum_{k\in S}p_{k,j}\left(n\right)p_{i,k}\left(m\right)\end{aligned}$$

□

Now let $P\left(1\right)=P=p_{i,j}$, $P\left(n\right)=p_{i,j}\left(n\right)$. Then this is just matrix multiplication: $P\left(n\right)=P^n$. To find $P\left(n\right)$ we can diagonalize the matrix. (or matrix mult + qpower!)

Example. Let $S=\{1,2\}$,

$$P=\begin{pmatrix}1-\alpha & \alpha\\ \beta & 1-\beta\end{pmatrix}$$

ans assume that $0<\alpha,\beta<1$ (non-trivial).
Then solve $|P-\kappa I|=0$, we get $\kappa=1$ or $\kappa=1-\alpha-\beta$. So

$$P^n=U^{-1}\begin{pmatrix}1^n & 0\\ 0 & \left(1-\alpha-\beta\right)^n\end{pmatrix}U$$

for some invertible U. Then

$$p_{1,1}\left(n\right)=A\cdot 1^n+B\left(1-\alpha-\beta\right)^n$$

for some A,B. We know that $p_{1,1}\left(1\right)=1-\alpha$, $p_{1,1}\left(0\right)=1$. Then we can solve for A and B and get

$$A=\frac{\beta}{\alpha+\beta},B=\frac{\alpha}{\alpha+\beta}$$

By symmetry we can get

$$P^n=\frac{1}{\alpha+\beta}\begin{pmatrix}\beta+\alpha\left(1-\alpha-\beta\right)^n & \alpha-\alpha\left(1-\alpha-\beta\right)^n\\ \beta-\beta\left(1-\alpha-\beta\right)^n & \alpha+\beta\left(1-\alpha-\beta\right)^n\end{pmatrix}$$

Another method is to use difference equations:

$$\begin{aligned} p_{1,1}(n+1) &= \mathbb{P}(X_{n+1}=1|X_0=1) \\ &= \mathbb{P}(X_{n+1}=1|X_n=1,X_0=1)\,\mathbb{P}(X_n=1|X_0=1)+ \\ &\mathbb{P}(X_{n+1}=1|X_n=2,X_0=1)\,\mathbb{P}(X_n=2|X_0=1) \\ &= (1-\alpha)\,P_{1,1}(n)+\beta p_{1,2}(n) \end{aligned}$$

which is a difference equation for the sequence for $(p_{1,1}(n))$ (note $p_{1,2}(n) = 1 - p_{1,1}(n)$) solved in the normal way, subject to boundary conditions.

The distributions of a Markov chain is somewhat related to linear algebra. Let λ be the initial distribution of X_0, i.e. $\lambda_i = \mathbb{P}(X_0 = i)$. Then

$$\mathbb{P}(X_1 = j) = \sum_i \lambda_i p_{i,j}$$

So the distribution of X_1 is λP, and similarly X_n has distribution λP^n.

3 Class Structure

We write "i leads to j", or $i \to j$, if there exists $n \geq 0$ s.t. $p_{i,j}(n) > 0$. Write $i \leftrightarrow j$ if $i \to j$ and $j \to i$, and say that i and j *communicate*.

Proposition. $\leftrightarrow$ is an equivalence relation.

Proof. • $i \leftrightarrow i$ since $p_{i,i}(0) = 1 > 0$.
• $i \leftrightarrow j \to j \leftrightarrow i$ is trivial.
• If $i \leftrightarrow j$ and $j \leftrightarrow i$, in particular $i \to j$ and $j \to k$. Then there exists m, k such that $p_{i,j}(m) > 0$ and $p_{j,k}(n) > 0$. Then

$$p_{i,k}(m+n) \geq p_{i,j}(m)\, p_{j,k}(n) > 0$$

By C-K equation (since each term in the sum is non-negative). So $i \to k$. Similarly $k \to i$. So $i \leftrightarrow k$. □

Definition. Now S, the set of all states, has equivalence classes under $\leftrightarrow$. We call them *communicating classes*, and define

$$C_i = \{j \in S : i \leftrightarrow j\}$$

The space S, or the chain X, is called *irreducible* if there exists a unique communicating class (which is S).

$C \subset S$ is called *closed* if

$$i \in C, i \to j \implies j \to C$$

if $\{i\}$ is closed, then i is called *absorbing*.

Proposition. A set C is closed if and only if $p_{i,j} = 0$ for all $i \in C, j \notin C$.

Proof. Suppose the above condition does not hold. Then $\exists i \in C, j \notin C$ with $p_{i,j} > 0$. But then C is not closed by definition since $i \to j$.
Now suppose the above condition hold. Let $i \in C, i \to j$. There exists several $k_0 = i, k_1, ..., k_n = j$ such that

$$p_{k_0,k_1} p_{k_1,k_2} ... p_{k_{n-1},k_n} > 0$$

which requires $i = k_0 \to k_1$, $k_1 \to k_2$, ... $k_{n-1} \to k_n = j$. Then $k_1 \in C, k_2 \in C, ... k_n = j \in C$. So C is closed. □

Example. Let $S = \{1, 2, ..., 6\}$, and

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Here $\{1, 2, 3\}$, $\{4\}$, $\{5, 6\}$ are communicating classes.

4 Recurrence and Transience

We write $\mathbb{P}\left(\cdot|X_0 = i\right) = \mathbb{P}_i\left(\cdot\right)$, and similarly for expectation.

Definition. The *first-passage time* of $j \in S$ is

$$T_j = \inf\left\{n \geq 1 : X_n = j\right\}$$

The *first-passage probabilities* are

$$f_{i,j}\left(n\right) = \mathbb{P}_i\left(T_j = n\right)$$

Definition. The state $i \in S$ is *recurrent* (or *persistent*) if $\mathbb{P}_i\left(T_i < \infty\right) = 1$, and *transient* otherwise.

Theorem. The state i is recurrent if and only if

$$\sum_n p_{ii}\left(n\right) \to \infty.$$

Proof. We have

$$\begin{aligned} p_{ij}\left(n\right) = \mathbb{P}_i\left(x_n = j\right) &= \sum_m \mathbb{P}_i\left(x_n = j|T_j = m\right)\mathbb{P}_i\left(T_j = m\right) \\ &= \sum_{m \leq n} \mathbb{P}_i\left(x_n = j|x_m = j\right)\mathbb{P}_i\left(T_j = m\right) \\ &= \sum_{m=1}^{n} f_{i,j}\left(m\right) p_{jj}\left(n - m\right) \end{aligned}$$

(which looks like a *convolution* of $f_{i,j}$ and p_{jj}).

Now consider generating sequences

$$\begin{aligned} F_{ij}\left(s\right) &= \sum_{n=0}^{\infty} f_{ij}\left(n\right) s^n \\ P_{ij}\left(s\right) &= \sum_{n=0}^{\infty} p_{ij}\left(n\right) s^n \end{aligned}$$

with $f_{ij}\left(0\right) = 0$,$p_{ij}\left(0\right) = \delta_{ij}$.

Then

$$\sum_{n \geq 1} p_{ij}\left(n\right) s^n = \sum_{n \geq 1} \sum_{m=1}^{n} f_{ij}\left(m\right) s^m p_{jj}\left(n - m\right) s^{n-m}$$

So by reversing the order of the sums,

$$P_{ij}\left(s\right) - \delta_{ij} = \sum_{m=0}^{\infty} f_{ij}\left(m\right) s^m \sum_{n=m}^{\infty} p_{jj}\left(n - m\right) s^{n-m} = F_{ij}\left(s\right) P_{jj}\left(s\right)$$

So we've derived

Theorem. $P_{ij}(s) = \delta_{ij} + F_{ij}(s) P_{jj}(s)$. $(1 < s \le 1$, since we need the series to converge)
When $i = j$,

$$P_{ii}(s) = \frac{1}{1 - F_{ii}(s)}$$

if $0 \le s < 1$.
Now let $s \to 1$, then

$$F_{ii}(s) \to \sum_n f_{ii}(n) = \mathbb{P}_i(T_i < \infty),$$

$$P_{ii}(s) \to \sum_n p_{ii}(n)$$

Therefore i is recurrent iff $F_{ij}(s) = 1$, i.e. $\sum_n p_{ii}(n) \to \infty$.

□

Theorem. Let C be a communicating class.
(a) For $i, j \in C$, either both of them are recurrent, or both are transient (i.e. recurrence is a class property).
(b) If $i \in C$ is recurrent, then C is closed (i.e. a recurrent communicating class is closed).

Proof. (a) Let $i \leftrightarrow j$. Then

$$p_{ii}(m + k + n) \ge p_{ij}(m) p_{jj}(k) p_{ji}(n)$$

Pick m s.t. $p_{ij}(m) > 0$, and n s.t. $p_{ji}(n) > 0$. Then

$$\sum_k p_{ii}(m + k + n) \ge \alpha \sum_k p_{jj}(k)$$

for $\alpha > 0$.
Then if j is recurrent, then $\sum_k p_{jj}(k) \to \infty$, and hence $\sum_k p_{ii}(k) \to \infty$, i.e. i is recurrent, and vice versa.

(b) Suppose C is not closed. So $\exists j \in C$, $k \notin C$ with $p_{jk} > 0$.
If i is recurrent, so is j by (a). Then

$$1 - \mathbb{P}_j(T_j < \infty) = \mathbb{P}_j(\text{no return to } j) \ge p_{jk}$$

Since $k \notin C$. However that implies $p_{jk} \le 1 - 1 = 0$. Contradiction. □

Proposition. Let $i, j \in S$. If j is transient, then $p_{ij}(n) \to 0$ as $n \to \infty$.

Proof. $P_{ij}(s) = \delta_{ij} + F_{ij}(s) P_{jj}(s)$ for $-1 < s < 1$.
Now let $i \ne j$, and $s \to 1$. Then

$$P_{ij}(1) = F_{ij}(1) P_{jj}(1)$$

Since j is transient, $F_{ij}(1) < \infty$ and $P_{jj}(1) < \infty$.
So $P_{ij}(n) < \infty$, and hence $p_{ij}(n) \to 0$ as $n \to \infty$.
The argument is similar when $i = j$. □

Theorem. If S is finite, then there exists at least one recurrent state. Therefore, if the chain is irreducible, every state is recurrent.

Proof. Suppose otherwise, that every j is transient. Then we have

$$1 = \sum_{j \in S} p_{ij}(n) \to 0$$

as $n \to \infty$. Contradiction. $\square$

5 Random walks on $\mathbb{Z}^d$ with $d \geq 1$

We consider random walks on

$$\mathbb{Z}^d = \{(x_1, ..., x_d) : x_i \in \mathbb{Z}\}$$

Define two points $x, y \in \mathbb{Z}^d$: $x = (x_1, ..., x_d)$, $y = (y_1, ..., y_d)$ to be *adjacent* if

$$\sum_{i=1}^{d} |x_i - y_i| = 1$$

The *Random walk* on $\mathbb{Z}^d$ is a Markov chain with state space $\mathbb{Z}^d$; a walker lives at X_n at time n, with

$$\mathbb{P}\left(X_{n+1} = y\right) | X_n = x, X_0 = x\left(0\right), ..., X_{n-1} = x\left(n-1\right) = \begin{cases} 0 & x, y \text{ are not adjacent} \\ \frac{1}{2d} & x, y \text{ are adjacent} \end{cases}$$

Clearly RW is irreducible and has an infinite state space.

Theorem. The random walk is recurrent if $d = 1, 2$, and transient if $d \geq 3$.

Proof. • $d = 1$: $p_{0,0}\left(2n\right) = \left(\frac{1}{2}\right)^{2n} \binom{2n}{n} = \left(\frac{1}{2}\right)^{2n} \frac{2n!}{(n!)^2} \sim \frac{1}{\sqrt{\pi n}}$ by Stirling formula. Hence $\sum_n p_{0,0}\left(n\right) \to \infty$, i.e. 0 is recurrent.

• $d = 2$: Suppose we walked m steps towards L/R, and $n - m$ steps towards U/D.

$$\begin{aligned} p_{0,0}\left(2n\right) &= \sum_{m=0}^{n} \left(\frac{1}{4}\right)^{2n} \frac{(2n)!}{m!m!\left(n-m\right)!\left(n-m\right)!} \\ &= \left(\frac{1}{4}\right)^{2n} \binom{2n}{n} \sum_{m=0}^{n} \binom{n}{m}^2 \\ &= \left(\frac{1}{4}\right)^{2n} \binom{2n}{n} \\ &\sim \frac{1}{\pi n} \end{aligned}$$

So $(0, 0)$ is recurrent.

• $d = 3$ (and similarly $d \geq 3$):

$$\begin{aligned} p_{0,0}\left(2n\right) &= \sum_{i+j+k=n} \left(\frac{1}{6}\right)^{2n} \frac{(2n)!}{\left(i!j!k!\right)^2} \\ &\leq \left(\frac{1}{2}\right)^{2n} \binom{2n}{n} \sum_{i+j+k=n} \left(\frac{n!}{3^n i!j!k!}\right)^2 \\ &\leq \left(\frac{1}{2}\right)^{2n} \binom{2n}{n} M_n \sum_{i+j+k=n} \frac{n!}{3^n i!k!l!} \end{aligned}$$

where

$$M_n = \max\left\{\frac{n!}{3^n i!j!k!} | i+j+k=n\right\}$$

The sum in the last line is 1, since each term in the sum is the probability of n balls goes in to 3 boxes, with i, j, k balls in each box.
We see that M_n is achieved when i, j, k are 'as equal as possible. Then

$$\begin{aligned} p_{0,0}(2n) &\leq \left(\frac{1}{2}\right)^2 n \binom{2n}{n} \frac{n!}{3^n \left(\lfloor n/3 \rfloor !\right)^3} \\ &\sim \frac{c}{n^{3/2}} \end{aligned}$$

But this sum now converges. So $(0,0,0)$ is transient. □

6 Hitting probabilities

6.1 Gambler's Ruin

What is the hitting probability for gambler's ruin, $h_i = h_i^{\{0\}}$?

$h_0 = 1$, $h_i = ph_{i+1} + qh_{i-1}$ for $i \geq 1$.

Then guess a solution $h_i = \theta^i$, so $\theta = q/p, 1$. So the general solution is

$$h_i = A + B\left(q/p\right)^i$$

for all i.

Since $h_0 = 1$, $A + B = 1$.

If $p < q$: since the h_i are probability, $B = 0$, $A = 1$. So $h_i = 1 - \left(q/p\right)^i$ for all i.

If $p > q$, since$h_i \geq 0$ for all i, we have $A \geq 0$. By minimality of (h_i), $A = 0$. So $h_i = \left(q/p\right)^i$.

When $p = q$: by the above arguments, $h_i \equiv 1$.

Extension: let $p_i = 1 - q_i \in (0, 1)$.

So $h_0 = 1$, $\left(p_i + q_i\right) h_i = p_i h_{i+1} + q_i h_{i-1}$. $p_i \left(h_{i+1} - h_i\right) = q_i \left(h_i - h_{i-1}\right)$.

Let $u_i = h_{i-1} - h_i$. Then $p_i u_{i+1} = q_i u_i$. So $u_{i+1} = \left(q_i/p_i\right) u_i$.

Therefore $u_{i+1}\gamma_i u_1$, where

$$\gamma_i = \frac{q_1 q_2 ... q_i}{p_1 p_2 ... p_i}$$

for $i \geq 1$, and $\gamma_0 = 1$. Then

$$u_1 + u_2 + ... + u_i = \left(h_0 - h_i\right)$$
$$h_i = 1 - \left(u_1 + ... + u_i\right) = 1 - u_1\left(\gamma_0 + \gamma_1 + ... + \gamma_{i-1}\right)$$

Let $S = \sum_{i=0}^{\infty} \gamma_i$. If $S = \infty$, since $h_1 \geq 0$ we have $u_1 = 0$, and hence $h_i \equiv 1$.

If $S < \infty$, u_1 is maximised when $1 - u_1 S = 0$, i.e. $u_1 = 1/S$.

7 Stopping times

Consider Markov chain X.

Definition. A random variable T taking values in $\{0, 1, 2, ..., \} \cup \{\infty\}$ is a *stopping time* (for X) if for $n \geq 0$, the event $\{T = n\}$ is given 'in terms of' X_0, X_1,...,X_n only.

Hitting times are stopping times: $\left\{H^A = n\right\} = \{X_n \in A\} \cap (\cap_{0 \leq m < n} \{x_m \notin A\})$.

$H^A + 1$ is a stopping times is a stopping time, $H^A - 1$ is not in general a stopping time.

Definition. (Strong Markov Property) Let X be a Markov chain with transition matrix P, and let T be a stopping time for X. Given $T < \infty$ and $X_T = i$, then $Y = (X_T, X_{T+1}, ...)$ is a Markov chain with transition matrix P and initial state $Y_0 = i$, and Y is independent of $(X_0, ..., X_{T-1})$.

Example. Consider a random walk with an absorbing wall at 0,with probability p going right and $q = 1 - p$ going left. Assume particle starts at 1, and let H be the hitting time of 0. What is the mass function and mean of H?

Let

$$G(s) = \mathbb{E}_1\left(s^H\right) = \sum_{n=0}^{\infty} s^n \mathbb{P}_1(H = n)$$

if $|s| < 1$.

By assuming $|s| < 1$ (and using Abel's lemma when needed) we include the possibility $\mathbb{P}_1(H = \infty) > 0$. Then

$$\begin{aligned}
G(s) &= \mathbb{E}_1\left(s^H\right) \\
&= \mathbb{E}_1\left(s^H | X_1 = 2\right) p + \mathbb{E}_1\left(s^H | X_1 = 0\right) q \\
&= p\mathbb{E}_1\left(s^{1+H_1+H_0}\right) + qs \\
&= psG(s)^2 \left(= \mathbb{E}_1\left(s^{H_1}\right)\mathbb{E}_1\left(s^{H_2}\right)\right) + qs
\end{aligned}$$

where H_i is the time to go from $i + 1$ to i.

So

$$G(s) = \frac{1 \pm \sqrt{1 - 4pqs^2}}{2ps}$$

for $|s| < 1$.

Since G is continuous, the $\pm$ sign is the same for all s. Since G has to remain regular at $s = 0$, the $\pm$ sign has to be $-$ for all s. So

$$G(s) = \frac{1 - \sqrt{1 - 4pqs^2}}{2ps}$$

So

$$\mathbb{P}_1(H = 2k - 1) = \frac{(2k-2)!}{k!\,(k-1)!} \frac{(pq)^k}{p}$$

where $k \geq 1$.

We can also get

$$\begin{aligned} P(H<\infty) &= \lim_{s\to 1} G(s) \\ &= \frac{1-\sqrt{1-4pq}}{2p} \\ &= \frac{1-|q-p|}{2p} \\ &= \begin{cases} 1 & p\leq q \\ q/p & p>q \end{cases} \end{aligned}$$

Now let $p \leq q$. We want to find $\mathbb{E}_1(H)$. Differentiate G,

$$\begin{aligned} G' &= pG^2 + 2psGG' + q \\ \implies G'(s) &= \frac{pG^2+q}{1-2psG} \end{aligned}$$

Take the limit $s \to 1$,

$$G'(s) \to \begin{cases} \infty & p=q \\ \frac{1}{q-p} & p<q \end{cases}$$

8 Classification of states

Definition. (a) The *mean recurrence time* of $i \in S$ is

$$\begin{aligned} \mu_i &= \mathbb{E}_i\left(T_i\right) \\ &= \begin{cases} \infty & i \text{ is transient} \\ \sum_{n \geq 1} n f_{i,i}\left(n\right) & i \text{ is recurrent} \end{cases} \end{aligned}$$

(b) We call i *null* if $\mu_i = \infty$, and *non-null* or *positive* if $\mu_i < \infty$.

(c) The *period* d_i of $i \in S$ is $d_i = \gcd\left\{n \geq 1 : p_{i,i}\left(n\right) > 0\right\}$.
We call i *aperiodic* if $d_i = 1$.

(d) A state $i \in S$ ie *ergodic* if it is aperiodic and non-null recurrent.

Theorem. If $i \leftrightarrow j$ then
(a) they have the same period;
(b) if one is recurrent, so is the other;
(c) if one is positive recurrent, so is the other;
(d) if one is ergodic, so is the other.