8. Statistical Equilibrium and Classification of States: Discrete Time Markov Chains

8.1 Review

8.2 Statistical Equilibrium

8.3 Two-State Markov Chain

8.4 Existence of $P(\infty)$

8.5 Classification of States

8.6 Terminology (Summary)

Discrete Time Markov Chains

8.1 Review

$\{X_n\}$ possible states $n = 0, 1, 2, \ldots$

Markov Property

$$P\{X_{n+1}|X_0, X_1, \ldots, X_n\} = P\{X_{n+1}|X_n\}$$

$p_{ij}(n) = P\{X_{n+1} = j|X_n = i\}$ 1-step probabilities

If $p_{ij}(n) = p_{ij}$ the process is termed Time Homogeneous

S = state space = $\{0, 1, 2, \ldots\}$

$\sum_{j \in S} p_{ij} = 1$, $p_{ij} \geq 0$

$P = (p_{ij})$ $i, j \in S$ Stochastic Matrix

$a_i = P\{X_0 = i\}$ $X_0 =$ initial state

$\{a_i\}$ and P completely determine the process

$$
\begin{aligned}
a_j^{(n)} &= P\{X_n = j\} = \text{ Prob. of being at } X_n = j \text{ in } n \text{ steps} \\
&= \textstyle\sum_{i \in S} P\{X_n = j | X_0 = i\} a_i \\
p_{ij}^{(n)} &= \text{Prob. of going from } i \to j \text{ in } n \text{ steps}
\end{aligned}
$$

<u>Chapman-Kolmogorov Equations</u>

for any $k (0 \le k \le n)$ $\boxed{p_{ij}^{(n)} = \sum_{i \in S} p_{ir}^{(k)} p_{rj}^{(n-k)}}$

or if $P^{(n)} = (p_{ij}^{(n)})$ $\boxed{P^{(n)} = P^{(k)} P^{(n-k)}}$

$\boxed{P^{(n)} = P^n}$ $\Rightarrow$ $\boxed{a^{(n)} = a P^n}$

where

$$
\begin{aligned}
a &= (a_1, a_2, \dots, a_m) \qquad (1 \times m) \\
a^{(n)} &= (a_1^{(n)}, a_2^{(n)}, \dots, a_m^{(n)}) \quad (1 \times m) \\
P &= (P_{ij}) \quad (m \times m)
\end{aligned}
$$

8.2 Statistical Equilibrium

Question: After a sufficiently long time does the system settle down into a condition of statistical equilibrium?

$$a^{(n)} = aP^n \quad a^{(n)} : 1 \times k, \quad a : 1 \times k, \quad P : k \times k$$

Define $\Pi = \lim_{n \to \infty} a^{(n)} = a \lim_{n \to \infty} P^n = aP^{(\infty)}$

In order to settle into statistical equilibrium $P^{(\infty)}$ must exist.

Ex.

$$P = \begin{bmatrix} .1 & .2 & .7 \\ .2 & .4 & .4 \\ .1 & .3 & .6 \end{bmatrix} \qquad a = \begin{bmatrix} .13 \\ .31 \\ .56 \end{bmatrix}$$

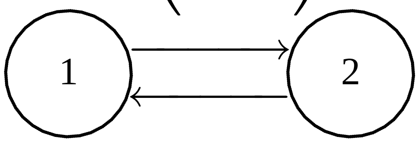
$$P^3 = \begin{bmatrix} .131 & .319 & .550 \\ .132 & .318 & .550 \\ .132 & .319 & .549 \end{bmatrix} \qquad a^{(3)} = \begin{bmatrix} .132 \\ .319 \\ .549 \end{bmatrix}$$

$$P^{(\infty)} = \begin{bmatrix} .132 & .319 & .549 \\ .132 & .319 & .549 \\ .132 & .319 & .549 \end{bmatrix} \qquad a^{(\infty)} = \begin{bmatrix} .132 \\ .319 \\ .549 \end{bmatrix}$$

Limit exists

Example:

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



(1) can only return to (1) in 2 steps

(2) can only return to (2) in 2 steps

$$P^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

$$P^3 = P, \;\; P^{2n} = I, \;\; P^{2n+1} = P, \;\; I + P = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Limit does not exist.

Consider the average

$$P^*(2n+1) = \frac{I + P + P^2 + \ldots + P^{2n+1}}{2n+2} = \frac{(n+1)}{2(n+1)}(I+P)$$

$$= \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{as } I + P = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$\therefore \qquad \lim_{n\to\infty} \frac{I + P + P^2 + \ldots + P^{2n+1}}{2n+2} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

In general if $P^{(\infty)}$ does not exist

$$P^*(\infty) = \lim_{n\to\infty} \frac{I + P + P^2 + \ldots + P^n}{n+1} \quad \text{does exist}$$

If $P^{(\infty)}$ does exist, $P^{(\infty)} = P^*(\infty)$.

$$P^*(n) = \frac{I + P + P^2 + \ldots + P^n}{n+1}$$

Returning to problem of equilibrium distribution

$$a^{(n)} = a^{(n-1)}P \qquad \Pi = \lim_{n\to\infty} a^{(n)}$$

Assume $P^{(\infty)}$ exists, then $\quad \Pi = \Pi P \quad$ or $\quad \Pi(I - P) = 0$
resulting in linear equations in Π. A solution exists if $|I - P| = 0$.

Recall $|P - \lambda I| = 0$ determines the eigenvalues.

Hence if $\lambda = 1$ is an eigenvalue $|I - P| = 0$.

Since P is stochastic $\sum_{j\in S} p_{ij} = 1$, all row sums are unity; i.e.

$$P\underline{1} = \underline{1} \qquad \underline{1}' = (1, 1, \ldots, 1)$$

The eigenvectors are defined by $Px = \lambda x$. In our case
$\lambda = 1, \;\; x = \underline{1}$ which shows $|P - I| = 0$

$$\Pi = \Pi P, \qquad \Pi = \lim_{n\to\infty} a^{(n)}.$$

Note however

$$P^{(n)} = P^n = P^{n-1}P$$

and as $n \to \infty$,

$$P^{(\infty)} = P^{(\infty)}P$$

$$P^{(\infty)}(I - P) = 0$$

Thus $\underline{1}\Pi = P^{(\infty)}$. Since $P^{(\infty)}$ does not involve $\underline{a}$ (initial conditions), the system in statistical equilibrium is independent of the initial conditions.

<u>Note:</u> $P(\infty) : k \times k \quad \underline{1}\Pi = \begin{pmatrix} \Pi \\ \Pi \\ \vdots \\ \Pi \end{pmatrix} : \ k \times k$

Spectral Decomposition

Suppose max eigenvalue is $\lambda_1 = 1$, all others are $|\lambda_i| < 1$ and λ_1 is of multiplicity one. The spectral decomposition is defined by being able to write P as:

$$P = \sum_{i=1}^{k} \lambda_i E_i = E_1 + \sum_{2}^{k} \lambda_i E_i$$

$$\begin{aligned}
P &= E_1 + \sum_{2}^{k} \lambda_i E_i, \quad E_i^2 = E_i \quad E_i E_j = 0 \quad i \neq j \\
P^n &= E_1 + \sum_{2}^{k} \lambda_i^n E_i \rightarrow E_1 \text{ as } n \rightarrow \infty \\
P^{\infty} &= \lim_{n \rightarrow \infty} P^n = E_1
\end{aligned}$$

E_1 can be found from left and right eigenvalues of P with $\lambda_1 = 1$.

$$
\begin{aligned}
Px &= \lambda_1 x = x \quad \text{(right eigenvector)} \quad x : k \times 1 \\
x &= \underline{1} \\
y'P &= \lambda_1 y' \quad \text{(left eigenvector)} \quad y : k \times 1
\end{aligned}
$$

Choose scale of y such that $\underline{1}'y = \sum_1^k y_i = 1$

$$
\begin{aligned}
E_1 &= xy' = \underline{1}y' \\
E_1^2 &= \underline{1}y'\underline{1}y' = \underline{1}y' = E_1
\end{aligned}
$$

Conclusion: If P has only a single eigenvalue equal to 1 and all others are $|\lambda_i| < 1 \Rightarrow P^\infty = E_1$ can easily be found.

8.3 Ex. Two State Markov Chains

$$S = \{0, 1\} \qquad P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$

We wish to write the spectral decomposition of P. $\Rightarrow$ Find the eigenvalues and eigenvectors

$$|P - \lambda I| = 0 \Rightarrow \begin{vmatrix} 1-\alpha-\lambda & \alpha \\ \beta & 1-\beta-\lambda \end{vmatrix} = 0$$

$$(1-\alpha-\lambda)(1-\beta-\lambda) - \alpha\beta = 0 \Rightarrow \lambda^2 - \lambda(2-\alpha-\beta) + (1-\alpha-\beta) = 0$$

$\Rightarrow$ $\boxed{\lambda_1 = 1, \ \lambda_2 = (1-\alpha-\beta)}$ roots are distinct provided $\alpha + \beta \neq 0$

$$Px = \lambda x \quad Px = x \quad \text{for } \lambda_1 = 1 \quad x = \underline{1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

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

i.e. $\underline{1}$ is the eigenvector as expected.

To obtain the right eigenvector corresponding to $\lambda_2 = (1-\alpha-\beta)$

$$P\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= \begin{pmatrix} (1-\alpha)x_1 + \alpha x_2 \\ \beta x_1 + (1-\beta)x_2 \end{pmatrix} = (1-\alpha-\beta)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\begin{aligned} (1-\alpha)x_1 + \alpha x_2 &= (1-\alpha-\beta)x_1 \Rightarrow \quad \beta x_1 = -\alpha x_2 \\ \beta x_1 + (1-\beta)x_2 &= (1-\alpha-\beta)x_2 \end{aligned}$$

Set $x_1 = \alpha$, then $x_2 = -\beta$

Define y as a (left) eigenvector of P.

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

If $y' = (y_1, y_2)$ we have $y'P = y'$ (left eigenvector associated with $\lambda = 1$)

$$(y_1 \;\; y_2) \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix} = (y_1 \;\; y_2)$$

$$\begin{bmatrix} y_1(1-\alpha) + y_2\beta & y_1\alpha + y_2(1-\beta) \end{bmatrix} = \begin{bmatrix} y_1 & y_2 \end{bmatrix}$$

$$\begin{matrix} y_1(1-\alpha) + y_2\beta = y_1 \\ y_1\alpha + y_2(1-\beta) = y_2 \end{matrix} \quad \Rightarrow \quad y_1 = y_2\frac{\beta}{\alpha},$$

Set $\quad y_2 = \alpha \Rightarrow \quad y_1 = \beta$

$$E = xy' = \begin{pmatrix} 1 \\ 1 \end{pmatrix} (\beta \quad \alpha) = \begin{pmatrix} \beta & \alpha \\ \beta & \alpha \end{pmatrix}$$

However since these are the limiting probabilities, the row sums must add to unity. We shall scale the eigenvector by $(\alpha + \beta)^{-1}$; i.e.

$$\Rightarrow \qquad E = \frac{1}{\alpha + \beta} \begin{bmatrix} \beta & \alpha \\ \beta & \alpha \end{bmatrix}$$

Check: $$E^2 = \frac{1}{(\alpha + \beta)^2} \begin{bmatrix} \beta^2 + \alpha\beta & \beta\alpha + \alpha^2 \\ \beta^2 + \alpha\beta & \beta\alpha + \alpha^2 \end{bmatrix} = \frac{1}{\alpha + \beta} \begin{bmatrix} \beta & \alpha \\ \beta & \alpha \end{bmatrix} = E$$

To obtain the left eigenvector corresponding to $\lambda = (1 - \alpha - \beta)$ we have

$$y'P = (1 - \alpha - \beta)y'$$

which can be written with $y' = (y_1 \;\; y_2)$

$$y'P = [y_1(1-\alpha) + \beta y_2 \quad y_1\alpha + y_2(1-\beta)] = (1-\alpha-\beta)[y_1 \quad y_2]$$

On solving $y_1\alpha = -\alpha y_2$ or $y_1 = -y_2$. We can take $y_1 = 1,\;\; y_2 = -1$. However it is necessary to divide by the scale factor $(\alpha + \beta)$. Therefore corresponding to $\lambda = (1 - \alpha - \beta)$ we have

$$E_2 = xy' = (\alpha+\beta)^{-1} \begin{bmatrix} \alpha \\ -\beta \end{bmatrix} [1 \;\; -1] = (\alpha+\beta)^{-1} \begin{bmatrix} \alpha & -\alpha \\ -\beta & \beta \end{bmatrix}.$$

We now can write

$$P = \lambda_1 E_1 + \lambda_2 E_2 = (\lambda + \beta)^{-1} \left\{ \begin{bmatrix} \beta & \alpha \\ \beta & \alpha \end{bmatrix} + (1 - \alpha - \beta) \begin{bmatrix} \alpha & -\alpha \\ -\beta & \beta \end{bmatrix} \right\}$$

and for P^n we have

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

$$\boxed{P^{(\infty)} = \frac{1}{\alpha + \beta} \begin{bmatrix} \beta & \alpha \\ \beta & \alpha \end{bmatrix}}$$

$P^{(\infty)}$ are equilibrium values

Note: To obtain $P^{(\infty)}$ directly it is only necessary to find the left and right eigenvectors associated with $\lambda = 1$.

8.4 Existence of $P^{(\infty)}$

Theorem   If $P^{(\infty)}$ exists it will always equal

$$P^*(\infty) = \lim_{n \to \infty} \frac{I + P + \ldots + P^n}{n+1}$$

Proof: Suppose $P^{(\infty)}$ exists; i.e. $P^{(\infty)} = E_1$ and P has only a single eigenvalue = 1.

$$P^*(n) = \frac{I + P + \ldots + P^n}{n+1} = \sum_{r=0}^{n} \frac{P^r}{n+1}$$

Suppose P $(m \times m)$.

$$P^r = \sum_{i=1}^{m} \lambda_i^r E_i = E_1 + \sum_{i=2}^{m} \lambda_i^r E_i \quad \lambda_i < 1, r \neq 0$$

$$P^*(n) = \frac{1}{n+1} \left\{ I + \sum_{r=1}^{n} \left[E_1 + \sum_{i=2}^{m} \lambda_i^r E_i \right] \right\}$$

$$= \frac{1}{n+1} \left\{ I + nE_1 + \sum_{i=2}^{m} E_i \sum_{r=1}^{n} \lambda_i^r \right\}$$

$$= \frac{1}{n+1} \left\{ I + nE_1 + \sum_{i=2}^{m} \frac{\lambda_i (1 - \lambda_i^n)}{1 - \lambda_i} E_i \right\}$$

as $n \to \infty, \quad P^*(\infty) = E_1$

$\Rightarrow \quad \boxed{P^*(\infty) = E_1 = P^{(\infty)}}$

8.5 Classification of States

Definition: A state j is accessible from state i if for some $n > 0$, $p_{ij}^{(n)} > 0$. We shall use the notation $i \rightarrow j$ to denote j is accessible from i.

Definition: If $i \rightarrow j$ and $j \rightarrow i$ the two states communicate; i.e. $p_{ij}^{(n)} > 0$, $p_{ji}^{(n\prime)} > 0$ for some n, n'.

Definition: A set $C \subset S$ is a communicating class if

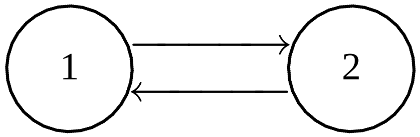
$$\text{(i)} \quad i \in C, \ \ j \in C \Rightarrow \ \ i \leftrightarrow \ j$$

$$\text{(ii)} \quad i \in C, \ \ i \leftrightarrow \ j \Rightarrow \ j \in C$$

Definition: A communicating class C is closed if $i \in C$ and $j \notin C \Rightarrow$ implies j is not accessible.

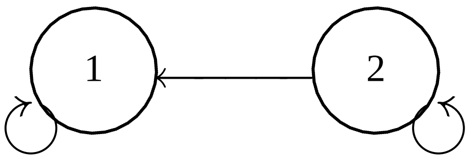
Def.: A Markov Chain is said to be irreducible if all states belong to a single closed communicating class. Otherwise it is called reducible.

Ex.



C={1, 2} is a closed communicating class.

Ex.



C={1} is a closed communicating class
C={2} is a communicating class which is not closed.

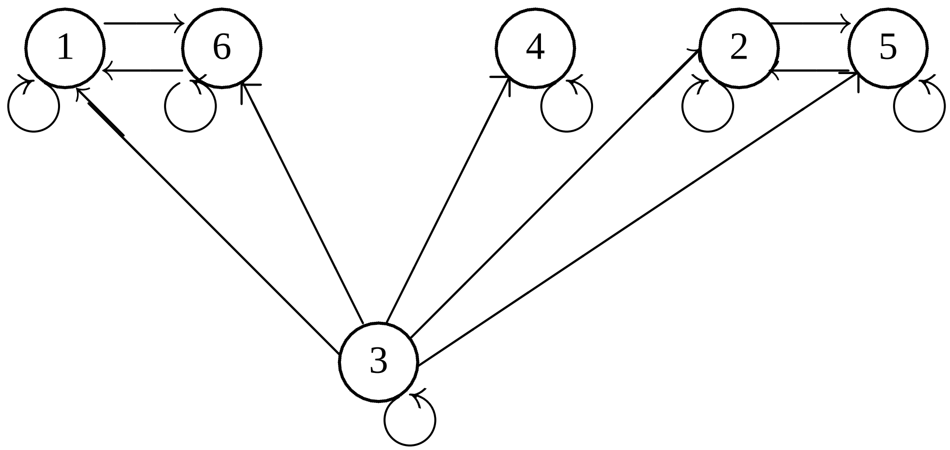
Note: All states communicate in an irreducible Markov Chain.

If P is reducible then by relabeling states we can can write

$$P = \begin{pmatrix} A & O \\ B & C \end{pmatrix}$$

Note that transitions from A to other states cannot happen.

Ex.

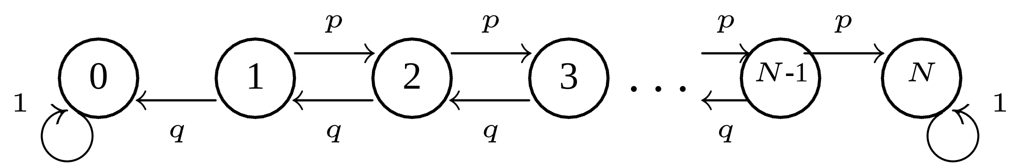


$C_1 = \{1, 6\}$, $C_2 = \{2, 5\}$, $C_3 = \{4\}$ are closed communicating classes.

$T = \{3\}$ is a communicating class which is not closed.

Ex. Random Walk with absorbing boundaries $S = \{0, 1, \ldots, N\}$

$$p_{00} = p_{NN} = 1, \quad p_{i,i+1} = p, \quad p_{i,i-1} = q, \quad p + q = 1$$



$C_1 = \{0\}$ and $C_2 = \{N\}$ are closed communicating classes

$T = \{1, 2, \ldots, N-1\}$ non-closed communicating class

Def. A state i is periodic with period d if d is the largest integer d such that $p_{ii}^{(n)} > 0$ where $n =$ integer multiple of d.

Def. A state i is aperiodic if $d = 1$.

Def. (Alternate): $T_i = \min\{n > 0 : X_n = i\}$

A state i is aperiodic with period d, if $d =$ largest integer such that

$$P\{T_i = n | X_0 = i\} > 0 \Rightarrow n \text{ is an integer multiple of } d$$

Ex. $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

If $X_0 = 1$, can only visit state 1 at times 2, 4, 6, ... Hence $d = 2$.

Since 1 $\longleftrightarrow$ 2, then state 2 is also periodic with $d = 2$.

8.6 Terminology Summary

Def. Accessible

State j is accessible from state C for some $n \geq 0$, $p_{ij}^{(n)} > 0 \quad (i \rightarrow j)$

Def. Communicate

States i and j communicate if each is accessible from the other ($i \leftrightarrow j$)

Def. Communicating Class

A set C is said to be a communicating class if

(i) $i \in C, j \in C \Rightarrow i \leftrightarrow j$

(ii) $i \in C, i \leftrightarrow j \Rightarrow j \in C$

Def. Closed Communicating Class

A communicating class is closed if $i \in C$ and $j \notin C$ implies j is not accessible from i.

Def. Irreducible

A Markov Chain is irreducible if all states belong to a single closed communicating class; i.e. all states in an irreducible chain communicate with each other.

Def. Reducible

A chain is reducible if by relabeling states, P can be written

$$P = \begin{pmatrix} A & O \\ B & C \end{pmatrix}$$

Def. Periodicity

A state is periodic with period d if d is largest integer such that

$$p_{ii}^{(n)} > 0 \Rightarrow n \text{ is integer multiple of } d$$

Def. Aperiodicity

A state i is aperiodic if $d = 1$.

Def. A state i is recurrent if starting initially from i $(X_0 = i)$ it returns to i with probability one $(f_i = 1)$.

Def. Transient

A state i is transient if $f_i < 1$.

Def. Positive and Null Recurrent

If m_i = mean time to return to state i $(X_0 = i)$, then state i is

positive recurrent if $m_i < \infty$

null recurrent if $m_i = \infty$

Def. Ergodicity

A state i is ergodic if it is aperiodic and positive recurrent.

Def. Absorbing State

A state i is absorbing if once entered cannot leave; i.e. closed set consisting of a single state.