CS70: Jean Walrand: Lecture 33.

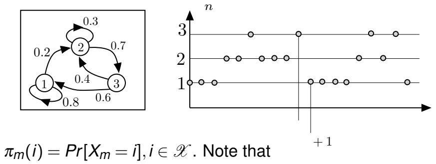
Markov Chains 2

1. Review
2. Distribution
3. Irreducibility
4. Convergence

Review

- Markov Chain:
  - Finite set $\mathscr{X}$; π_0; $P = \{P(i,j), i,j \in \mathscr{X}\}$;
  - $Pr[X_0 = i] = \pi_0(i), i \in \mathscr{X}$
  - $Pr[X_{n+1} = j \mid X_0, \ldots, X_n = i] = P(i,j), i,j \in \mathscr{X}, n \geq 0$.
  - Note:
    $Pr[X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n] = \pi_0(i_0)P(i_0,i_1)\cdots P(i_{n-1},i_n)$.
- First Passage Time:
  - $A \cap B = \emptyset; \beta(i) = E[T_A | X_0 = i]; \alpha(i) = P[T_A < T_B | X_0 = i]$
  - $\beta(i) = 1 + \sum_j P(i,j)\beta(j); \alpha(i) = \sum_j P(i,j)\alpha(j)$.

Distribution of X_n



Let $\pi_m(i) = Pr[X_m = i], i \in \mathscr{X}$. Note that

$$\begin{aligned} Pr[X_{m+1} = j] &= \sum_i Pr[X_{m+1} = j, X_m = i] \\ &= \sum_i Pr[X_m = i]Pr[X_{m+1} = j \mid X_m = i] \\ &= \sum_i \pi_m(i)P(i,j). \end{aligned}$$

Hence,

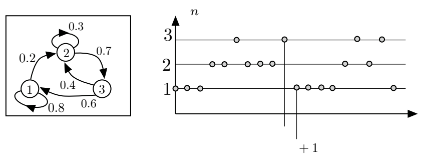
$$\pi_{m+1}(j) = \sum_i \pi_m(i)P(i,j), \forall j \in \mathscr{X}.$$

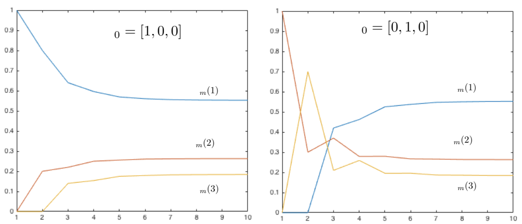
With π_m, π_{m+1} as a row vectors, these identities are written as $\pi_{m+1} = \pi_m P$.

Thus, $\pi_1 = \pi_0 P$, $\pi_2 = \pi_1 P = \pi_0 PP = \pi_0 P^2, \ldots$. Hence,

$$\pi_n = \pi_0 P^n, n \geq 0.$$

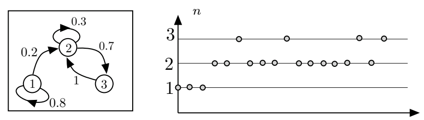
Distribution of X_n

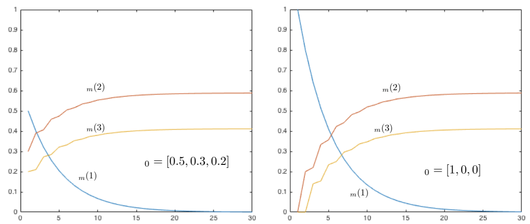




As m increases, π_m converges to a vector that does not depend on π_0.

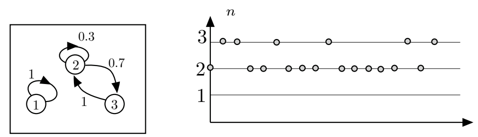
Distribution of X_n

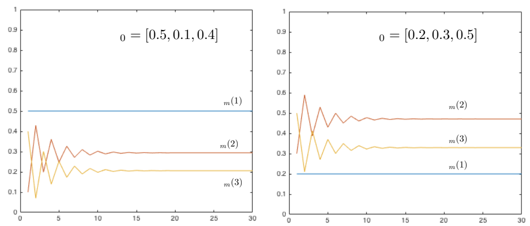




As m increases, π_m converges to a vector that does not depend on π_0.

Distribution of X_n





As m increases, π_m converges to a vector that depends on π_0 (obviously, since $\pi_m(1) = \pi_0(1), \forall m$).

Balance Equations

Question: Is there some π_0 such that $\pi_m = \pi_0, \forall m$?

Definition A distribution π_0 such that $\pi_m = \pi_0, \forall m$ is said to be an invariant distribution.

Theorem A distribution π_0 is invariant iff $\pi_0 P = \pi_0$. These equations are called the balance equations.

Proof: $\pi_n = \pi_0 P^n$, so that $\pi_n = \pi_0, \forall n$ iff $\pi_0 P = \pi_0$. □

Thus, if π_0 is invariant, the distribution of X_n is always the same as that of X_0.

Of course, this does not mean that X_n does not move. It means that the probability that it leaves a state i is equal to the probability that it enters state i.

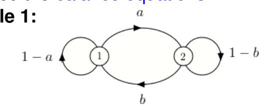
The balance equations say that $\sum_j \pi(j)P(j,i) = \pi(i)$.
That is,

$$\sum_{j \neq i} \pi(j)P(j,i) = \pi(i)(1 - P(i,i)) = \pi(i)\sum_{j \neq i} P(i,j).$$

Thus, $Pr[\text{enter } i] = Pr[\text{leave } i]$.

Balance Equations

Theorem A distribution π_0 is invariant iff $\pi_0 P = \pi_0$. These equations are called the balance equations.
Example 1:



$$P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}$$

$$\begin{aligned}
\pi P = \pi \quad &\Leftrightarrow \quad [\pi(1), \pi(2)] \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} = [\pi(1), \pi(2)] \\
&\Leftrightarrow \quad \pi(1)(1-a) + \pi(2)b = \pi(1) \text{ and } \pi(1)a + \pi(2)(1-b) = \pi(2) \\
&\Leftrightarrow \quad \pi(1)a = \pi(2)b.
\end{aligned}$$

These equations are redundant! We have to add an equation: $\pi(1) + \pi(2) = 1$. Then we find

$$\pi = [\frac{b}{a+b}, \frac{a}{a+b}].$$

Balance Equations

Theorem A distribution π_0 is invariant iff $\pi_0 P = \pi_0$. These equations are called the balance equations.
Example 2:

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

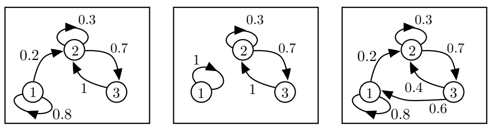
$$\pi P = \pi \Leftrightarrow [\pi(1), \pi(2)] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = [\pi(1), \pi(2)] \Leftrightarrow \pi(1) = \pi(1) \text{ and } \pi(2) = \pi(2).$$

Every distribution is invariant for this Markov chain. This is obvious, since $X_n = X_0$ for all n. Hence, $Pr[X_n = i] = Pr[X_0 = i], \forall (i,n)$.

Irreducibility

Definition A Markov chain is irreducible if it can go from every state i to every state j (possibly in multiple steps).

Examples:



[A] [B] [C]

[A] is not irreducible. It cannot go from (2) to (1).

[B] is not irreducible. It cannot go from (2) to (1).

[C] is irreducible. It can go from every i to every j.

If you consider the graph with arrows when $P(i,j) > 0$, irreducible means that there is a single connected component.

Existence and uniqueness of Invariant Distribution

Theorem A finite irreducible Markov chain has one and only one invariant distribution.

That is, there is a unique positive vector $\pi = [\pi(1), \ldots, \pi(K)]$ such that $\pi P = \pi$ and $\sum_k \pi(k) = 1$.

Proof: See EE126, or lecture note 24. (We will not expect you to understand this proof.)
Note: We know already that some irreducible Markov chains have multiple invariant distributions.

Fact: If a Markov chain has two different invariant distributions π and ν, then it has infinitely many invariant distributions. Indeed, $p\pi + (1-p)\nu$ is then invariant since

$$[p\pi + (1-p)\nu]P = p\pi P + (1-p)\nu P = p\pi + (1-p)\nu.$$

Long Term Fraction of Time in States

Theorem Let X_n be an irreducible Markov chain with invariant distribution π.

Then, for all i,

$$\frac{1}{n}\sum_{m=0}^{n-1} 1\{X_m = i\} \to \pi(i), \text{ as } n \to \infty.$$

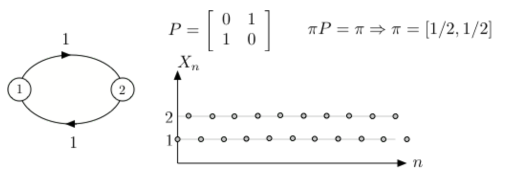
The left-hand side is the fraction of time that $X_m = i$ during steps $0, 1, \ldots, n-1$. Thus, this fraction of time approaches $\pi(i)$.

Proof: See EE126. Lecture note 24 gives a plausibility argument. □

Long Term Fraction of Time in States

Theorem Let X_n be an irreducible Markov chain with invariant distribution π. Then, for all i, $\frac{1}{n}\sum_{m=0}^{n-1} 1\{X_m = i\} \to \pi(i)$, as $n \to \infty$.

Example 1:

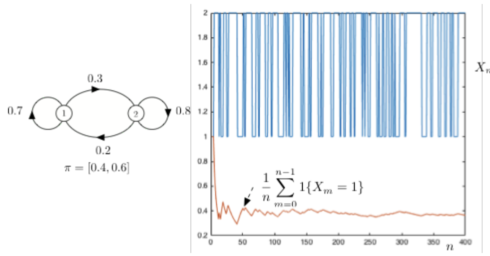


The fraction of time in state 1 converges to $1/2$, which is $\pi(1)$.

Long Term Fraction of Time in States

Theorem Let X_n be an irreducible Markov chain with invariant distribution π. Then, for all i, $\frac{1}{n}\sum_{m=0}^{n-1} 1\{X_m = i\} \to \pi(i)$, as $n \to \infty$.

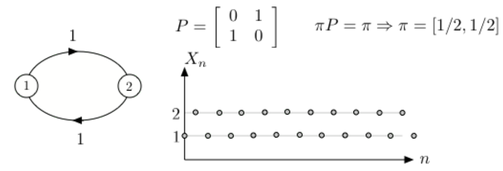
Example 2:



Convergence to Invariant Distribution

Question: Assume that the MC is irreducible. Does π_n approach the unique invariant distribution π?

Answer: Not necessarily. Here is an example:



Assume $X_0 = 1$. Then $X_1 = 2, X_2 = 1, X_3 = 2, \ldots$.

Thus, if $\pi_0 = [1, 0]$, $\pi_1 = [0, 1], \pi_2 = [1, 0], \pi_3 = [0, 1]$, etc.

Hence, π_n does not converge to $\pi = [1/2, 1/2]$.

Periodicity

Theorem Assume that the MC is irreducible. Then

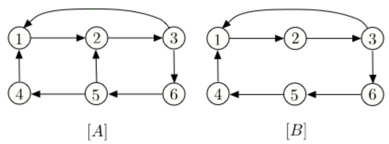
$$d(i) := \text{g.c.d.}\{n > 0 \mid Pr[X_n = i \mid X_0 = i] > 0\}$$

has the same value for all states i.

Proof: See Lecture notes 24. □

Definition If $d(i) = 1$, the Markov chain is said to be aperiodic. Otherwise, it is periodic with period $d(i)$.

Example



[A]: $\{n > 0 \mid Pr[X_n = 1 | X_0 = 1] > 0\} = \{3, 6, 7, 9, 11, \ldots\} \Rightarrow d(1) = 1$.
$\{n > 0 \mid Pr[X_n = 2 | X_0 = 2] > 0\} = \{3, 4, \ldots\} \Rightarrow d(2) = 1$.

[B]: $\{n > 0 \mid Pr[X_n = 1 | X_0 = 1] > 0\} = \{3, 6, 9, \ldots\} \Rightarrow d(i) = 3$.
$\{n > 0 \mid Pr[X_n = 5 | X_0 = 5] > 0\} = \{6, 9, \ldots\} \Rightarrow d(5) = 3$.

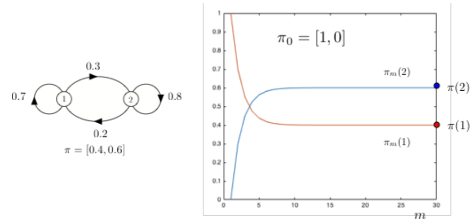
Convergence of π_n

Theorem Let X_n be an irreducible and aperiodic Markov chain with invariant distribution π. Then, for all $i \in \mathscr{X}$,

$$\pi_n(i) \to \pi(i), \text{ as } n \to \infty.$$

Proof See EE126, or Lecture notes 24. □

Example



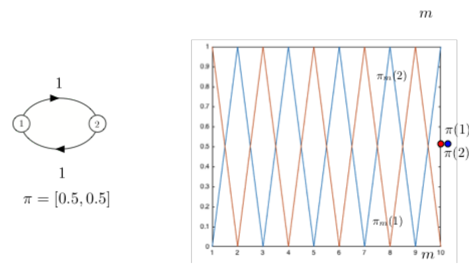
Convergence of π_n

Theorem Let X_n be an irreducible and aperiodic Markov chain with invariant distribution π. Then, for all $i \in \mathscr{X}$,

$$\pi_n(i) \to \pi(i), \text{ as } n \to \infty.$$

Proof See EE126, or Lecture notes 24. □

Example



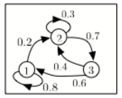
Convergence of π_n

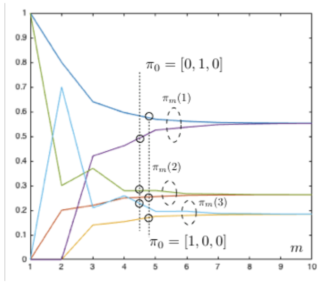
Theorem Let X_n be an irreducible and aperiodic Markov chain with invariant distribution π. Then, for all $i \in \mathscr{X}$,

$$\pi_n(i) \to \pi(i), \text{ as } n \to \infty.$$

Proof See EE126, or Lecture notes 24. $\square$

Example





Calculating π

Let P be irreducible. How do we find π?

Example: $P = \begin{bmatrix} 0.8 & 0.2 & 0 \\ 0 & 0.3 & 0.7 \\ 0.6 & 0.4 & 0 \end{bmatrix}$.

One has $\pi P = \pi$, i.e., $\pi[P - I] = \mathbf{0}$ where I is the identity matrix:

$$\pi \begin{bmatrix} 0.8-1 & 0.2 & 0 \\ 0 & 0.3-1 & 0.7 \\ 0.6 & 0.4 & 0-1 \end{bmatrix} = [0,0,0].$$

However, the sum of the columns of $P - I$ is $\mathbf{0}$. This shows that these equations are redundant: If all but the last one hold, so does the last one. Let us replace the last equation by $\pi \mathbf{1} = 1$, i.e., $\sum_j \pi(j) = 1$:

$$\pi \begin{bmatrix} 0.8-1 & 0.2 & 1 \\ 0 & 0.3-1 & 1 \\ 0.6 & 0.4 & 1 \end{bmatrix} = [0,0,1].$$

Hence,

$$\pi = [0,0,1] \begin{bmatrix} 0.8-1 & 0.2 & 1 \\ 0 & 0.3-1 & 1 \\ 0.6 & 0.4 & 1 \end{bmatrix}^{-1} \approx [0.55, 0.26, 0.19]$$

Summary

$\boxed{\text{Markov Chains}}$

- Markov Chain: $Pr[X_{n+1} = j | X_0, \dots, X_n = i] = P(i,j)$
- FSE: $\beta(i) = 1 + \sum_j P(i,j)\beta(j); \alpha(i) = \sum_j P(i,j)\alpha(j)$.
- $\pi_n = \pi_0 P^n$
- π is invariant iff $\pi P = \pi$
- Irreducible $\Rightarrow$ one and only one invariant distribution π
- Irreducible $\Rightarrow$ fraction of time in state i approaches $\pi(i)$
- Irreducible + Aperiodic $\Rightarrow \pi_n \to \pi$.
- Calculating π: One finds $\pi = [0, 0 \dots, 1]Q^{-1}$ where $Q = \cdots$.