CS70: Lecture 36.

Markov Chains

CS70: Lecture 36.

Markov Chains

CS70: Lecture 36.

Markov Chains

1. Markov Process: Motivation, Definition
2. Examples
3. Invariant Distribution of Markov Chains: Balance Equations

From Random Variables to Random Processes

What is a random **process?**

From Random Variables to Random Processes

What is a random **process?**

$\Rightarrow$ Probabilistic description for a **sequence** of Random Variables $\Rightarrow$ usually associated with

From Random Variables to Random Processes

What is a random **process?**

$\Rightarrow$ Probabilistic description for a **sequence** of Random Variables $\Rightarrow$ usually associated with **time.**

From Random Variables to Random Processes

What is a random **process?**

$\Rightarrow$ Probabilistic description for a **sequence** of Random Variables $\Rightarrow$ usually associated with **time.**

Example 1: No. of students in my Office Hours (OH) at time t (5-minute intervals)

From Random Variables to Random Processes

What is a random **process?**

$\Rightarrow$ Probabilistic description for a **sequence** of Random Variables $\Rightarrow$ usually associated with **time.**

Example 1: No. of students in my Office Hours (OH) at time t (5-minute intervals)

Example 2: No. of dollars in my wallet at the end of a day

From Random Variables to Random Processes

What is a random **process?**

$\Rightarrow$ Probabilistic description for a **sequence** of Random Variables $\Rightarrow$ usually associated with **time.**

Example 1: No. of students in my Office Hours (OH) at time t (5-minute intervals)

Example 2: No. of dollars in my wallet at the end of a day

$X_{11/29/17} = \$17$

From Random Variables to Random Processes

What is a random **process?**

$\Rightarrow$ Probabilistic description for a **sequence** of Random Variables $\Rightarrow$ usually associated with **time.**

Example 1: No. of students in my Office Hours (OH) at time t (5-minute intervals)

Example 2: No. of dollars in my wallet at the end of a day

$X_{11/29/17} = \$17$

$X_{11/30/17} = \$7$ with probability 0.5

From Random Variables to Random Processes

What is a random **process?**

$\Rightarrow$ Probabilistic description for a **sequence** of Random Variables $\Rightarrow$ usually associated with **time.**

Example 1: No. of students in my Office Hours (OH) at time t (5-minute intervals)

Example 2: No. of dollars in my wallet at the end of a day

$X_{11/29/17} = \$17$

$X_{11/30/17} = \$7$ with probability 0.5 and $= \$13$ with probability 0.5

From Random Variables to Random Processes

What is a random **process?**

$\Rightarrow$ Probabilistic description for a **sequence** of Random Variables $\Rightarrow$ usually associated with **time.**

Example 1: No. of students in my Office Hours (OH) at time t (5-minute intervals)

Example 2: No. of dollars in my wallet at the end of a day

$X_{11/29/17} = \$17$

$X_{11/30/17} = \$7$ with probability 0.5 and $= \$13$ with probability 0.5

Example 3: No. of students enrolled in CS70:

From Random Variables to Random Processes

What is a random **process?**

$\Rightarrow$ Probabilistic description for a **sequence** of Random Variables $\Rightarrow$ usually associated with **time.**

Example 1: No. of students in my Office Hours (OH) at time t (5-minute intervals)

Example 2: No. of dollars in my wallet at the end of a day

$X_{11/29/17} = \$17$

$X_{11/30/17} = \$7$ with probability 0.5 and $= \$13$ with probability 0.5

Example 3: No. of students enrolled in CS70:

Sept. 1: 800;

From Random Variables to Random Processes

What is a random **process?**

$\Rightarrow$ Probabilistic description for a **sequence** of Random Variables $\Rightarrow$ usually associated with **time.**

Example 1: No. of students in my Office Hours (OH) at time t (5-minute intervals)

Example 2: No. of dollars in my wallet at the end of a day

$X_{11/29/17} = \$17$

$X_{11/30/17} = \$7$ with probability 0.5 and $= \$13$ with probability 0.5

Example 3: No. of students enrolled in CS70:

Sept. 1: 800; Oct. 1: 850;

From Random Variables to Random Processes

What is a random **process?**

$\Rightarrow$ Probabilistic description for a **sequence** of Random Variables $\Rightarrow$ usually associated with **time.**

Example 1: No. of students in my Office Hours (OH) at time t (5-minute intervals)

Example 2: No. of dollars in my wallet at the end of a day

$X_{11/29/17} = \$17$

$X_{11/30/17} = \$7$ with probability 0.5 and $= \$13$ with probability 0.5

Example 3: No. of students enrolled in CS70:

Sept. 1: 800; Oct. 1: 850; Nov. 1: 750;

From Random Variables to Random Processes

What is a random **process?**

$\Rightarrow$ Probabilistic description for a **sequence** of Random Variables $\Rightarrow$ usually associated with **time.**

Example 1: No. of students in my Office Hours (OH) at time *t* (5-minute intervals)

Example 2: No. of dollars in my wallet at the end of a day

$X_{11/29/17} = \$17$

$X_{11/30/17} = \$7$ with probability 0.5 and $= \$13$ with probability 0.5

Example 3: No. of students enrolled in CS70:

Sept. 1: 800; Oct. 1: 850; Nov. 1: 750; Dec. 1: 737;

Random Process

In general, one can describe a **random process** by describing the **joint distribution** of $(X_{t_1}, X_{t_2}, \ldots, X_{t_i})\ \forall i \Rightarrow$

Random Process

In general, one can describe a **random process** by describing the **joint distribution** of $(X_{t_1}, X_{t_2}, \ldots, X_{t_i})\ \forall i \Rightarrow$ *not tractable* **.**

Random Process

In general, one can describe a **random process** by describing the **joint distribution** of $(X_{t_1}, X_{t_2}, \ldots, X_{t_i})\ \forall i \Rightarrow$ *not tractable* **.**

Markov Process: We make the simplifying assumption:

Random Process

In general, one can describe a **random process** by describing the **joint distribution** of $(X_{t_1}, X_{t_2}, \ldots, X_{t_i})\ \forall i \Rightarrow$ *not tractable* .

Markov Process: We make the simplifying assumption:

"Given the present, the future is decoupled from the past."

Random Process

In general, one can describe a **random process** by describing the **joint distribution** of $(X_{t_1}, X_{t_2}, \ldots, X_{t_i})\ \forall i \Rightarrow$ *not tractable* .

Markov Process: We make the simplifying assumption:

"Given the present, the future is decoupled from the past."

Example: Suppose you need to get to an 8 a.m. class, and you need to take a 7:30 a.m. bus from near your house to make it on time to class.

Random Process

In general, one can describe a **random process** by describing the **joint distribution** of $(X_{t_1}, X_{t_2}, \ldots, X_{t_i})\ \forall i \Rightarrow$ *not tractable* .

Markov Process: We make the simplifying assumption:

"Given the present, the future is decoupled from the past."

Example: Suppose you need to get to an 8 a.m. class, and you need to take a 7:30 a.m. bus from near your house to make it on time to class.

Pr[You get to your 8 a.m. class on time | You catch the 7:30 bus,

Random Process

In general, one can describe a **random process** by describing the **joint distribution** of $(X_{t_1}, X_{t_2}, \ldots, X_{t_i})\ \forall i \Rightarrow$ *not tractable* .

Markov Process: We make the simplifying assumption:

"Given the present, the future is decoupled from the past."

Example: Suppose you need to get to an 8 a.m. class, and you need to take a 7:30 a.m. bus from near your house to make it on time to class.

Pr[You get to your 8 a.m. class on time | You catch the 7:30 bus, You wake up at 6 a.m.,

Random Process

In general, one can describe a **random process** by describing the **joint distribution** of $(X_{t_1}, X_{t_2}, \ldots, X_{t_i})\ \forall i \Rightarrow$ *not tractable* .

Markov Process: We make the simplifying assumption:

“Given the present, the future is decoupled from the past.”

Example: Suppose you need to get to an 8 a.m. class, and you need to take a 7:30 a.m. bus from near your house to make it on time to class.

Pr[You get to your 8 a.m. class on time | You catch the 7:30 bus, You wake up at 6 a.m., You eat breakfast at 7 a.m.]

Random Process

In general, one can describe a **random process** by describing the **joint distribution** of $(X_{t_1}, X_{t_2}, \ldots, X_{t_i})\ \forall i \Rightarrow$ *not tractable* .

Markov Process: We make the simplifying assumption:

"Given the present, the future is decoupled from the past."

Example: Suppose you need to get to an 8 a.m. class, and you need to take a 7:30 a.m. bus from near your house to make it on time to class.

Pr[You get to your 8 a.m. class on time | You catch the 7:30 bus, You wake up at 6 a.m., You eat breakfast at 7 a.m.]

= *Pr* [You get to your 8 a.m. class on time | You catch the 7:30 bus].

Random Process

In general, one can describe a **random process** by describing the **joint distribution** of $(X_{t_1}, X_{t_2}, \ldots, X_{t_i})$ $\forall i \Rightarrow$ *not tractable* .

Markov Process: We make the simplifying assumption:

"Given the present, the future is decoupled from the past."

Example: Suppose you need to get to an 8 a.m. class, and you need to take a 7:30 a.m. bus from near your house to make it on time to class.

Pr[You get to your 8 a.m. class on time | You catch the 7:30 bus, You wake up at 6 a.m., You eat breakfast at 7 a.m.]

= *Pr* [You get to your 8 a.m. class on time | You catch the 7:30 bus].

This is an example of the Markov property:
$P[X_{n+1} = x_{n+1} | X_n = x_n, X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \ldots]$

Random Process

In general, one can describe a **random process** by describing the **joint distribution** of $(X_{t_1}, X_{t_2}, \ldots, X_{t_i})\ \forall i \Rightarrow$ *not tractable* .

Markov Process: We make the simplifying assumption:

"Given the present, the future is decoupled from the past."

Example: Suppose you need to get to an 8 a.m. class, and you need to take a 7:30 a.m. bus from near your house to make it on time to class.

Pr[You get to your 8 a.m. class on time | You catch the 7:30 bus, You wake up at 6 a.m., You eat breakfast at 7 a.m.]

= *Pr* [You get to your 8 a.m. class on time | You catch the 7:30 bus].

This is an example of the Markov property:
$P[X_{n+1} = x_{n+1} | X_n = x_n, X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \ldots]$

$= P[X_{n+1} = x_{n+1} | X_n = x_n]$

Example: My Office Hours (OH)

- When nobody is in my OH at time n,

Example: My Office Hours (OH)

- When nobody is in my OH at time n, then at time (n+1),

Example: My Office Hours (OH)

- When nobody is in my OH at time n, then at time (n+1), there will be either 1 student w.p. 0.2 or 0 student w.p. 0.8

Example: My Office Hours (OH)

- When nobody is in my OH at time n, then at time (n+1), there will be either 1 student w.p. 0.2 or 0 student w.p. 0.8
- When 1 person is in my OH at time n,

Example: My Office Hours (OH)

- When nobody is in my OH at time n, then at time (n+1), there will be either 1 student w.p. 0.2 or 0 student w.p. 0.8
- When 1 person is in my OH at time n, then at time (n+1),

Example: My Office Hours (OH)

- When nobody is in my OH at time n, then at time (n+1), there will be either 1 student w.p. 0.2 or 0 student w.p. 0.8
- When 1 person is in my OH at time n, then at time (n+1), there will be either 1 student w.p. 0.3 or 2 students w.p. 0.7

Example: My Office Hours (OH)

- When nobody is in my OH at time n, then at time (n+1), there will be either 1 student w.p. 0.2 or 0 student w.p. 0.8
- When 1 person is in my OH at time n, then at time (n+1), there will be either 1 student w.p. 0.3 or 2 students w.p. 0.7
- When 2 people are in my OH at time n,

Example: My Office Hours (OH)

- When nobody is in my OH at time n, then at time (n+1), there will be either 1 student w.p. 0.2 or 0 student w.p. 0.8
- When 1 person is in my OH at time n, then at time (n+1), there will be either 1 student w.p. 0.3 or 2 students w.p. 0.7
- When 2 people are in my OH at time n, then at time (n+1),

Example: My Office Hours (OH)

- When nobody is in my OH at time n, then at time (n+1), there will be either 1 student w.p. 0.2 or 0 student w.p. 0.8
- When 1 person is in my OH at time n, then at time (n+1), there will be either 1 student w.p. 0.3 or 2 students w.p. 0.7
- When 2 people are in my OH at time n, then at time (n+1), there will be either 0 student w.p. 0.6 or 1 student w.p. 0.4

Example: My Office Hours (OH)

- When nobody is in my OH at time n, then at time (n+1), there will be either 1 student w.p. 0.2 or 0 student w.p. 0.8
- When 1 person is in my OH at time n, then at time (n+1), there will be either 1 student w.p. 0.3 or 2 students w.p. 0.7
- When 2 people are in my OH at time n, then at time (n+1), there will be either 0 student w.p. 0.6 or 1 student w.p. 0.4

Questions of interest:

Example: My Office Hours (OH)

- When nobody is in my OH at time n, then at time $(n+1)$, there will be either 1 student w.p. 0.2 or 0 student w.p. 0.8
- When 1 person is in my OH at time n, then at time $(n+1)$, there will be either 1 student w.p. 0.3 or 2 students w.p. 0.7
- When 2 people are in my OH at time n, then at time $(n+1)$, there will be either 0 student w.p. 0.6 or 1 student w.p. 0.4

Questions of interest:

1. How many students do I have in my OH on average?

Example: My Office Hours (OH)

- When nobody is in my OH at time n, then at time $(n+1)$, there will be either 1 student w.p. 0.2 or 0 student w.p. 0.8
- When 1 person is in my OH at time n, then at time $(n+1)$, there will be either 1 student w.p. 0.3 or 2 students w.p. 0.7
- When 2 people are in my OH at time n, then at time $(n+1)$, there will be either 0 student w.p. 0.6 or 1 student w.p. 0.4

Questions of interest:

1. How many students do I have in my OH on average?
2. If I start my OH at time 0, with 0 students, what is the probability that I have 2 students in my OH at time 10?

Example: My Office Hours (OH)

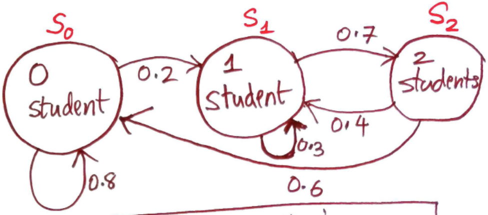
- When nobody is in my OH at time n, then at time (n+1), there will be either 1 student w.p. 0.2 or 0 student w.p. 0.8
- When 1 person is in my OH at time n, then at time (n+1), there will be either 1 student w.p. 0.3 or 2 students w.p. 0.7
- When 2 people are in my OH at time n, then at time (n+1), there will be either 0 student w.p. 0.6 or 1 student w.p. 0.4

Questions of interest:

1. How many students do I have in my OH on average?
2. If I start my OH at time 0, with 0 students, what is the probability that I have 2 students in my OH at time 10?

These questions require the study of **Markov Chains**!

State Transition Diagram and Matrix



State Transition Diagram

State Sequence: $X_0, X_1, X_2, \ldots$,
where $X_i \in \{S_0, S_1, S_2\}$ $\forall i = 1, 2, \ldots$

From \ To	S_0	S_1	S_2
S_0	0.8	0.2	0
S_1	0	0.3	0.7
S_2	0.6	0.4	0

- Row-Sums ≡ 1 ∀ Rows
- $$P = \begin{bmatrix} 0.8 & 0.2 & 0 \\ 0 & 0.3 & 0.7 \\ 0.6 & 0.4 & 0 \end{bmatrix}$$

state-transition matrix

$$P = \begin{bmatrix} P(0,0) & P(0,1) & P(0,2) \\ P(1,0) & P(1,1) & P(1,2) \\ P(2,0) & P(2,1) & P(2,2) \end{bmatrix}$$

$P(2,1) \rightarrow Pr[S_2 \rightarrow S_1]$

$$P(i,j) \equiv Pr[S_i \longrightarrow S_j]$$

$$\underline{\pi}_0 = [\pi_0(0) \quad \pi_0(1) \quad \pi_0(2)]$$

↑ Row-vector

$$= [P(X_0 = S_0) \quad P(X_0 = S_1) \quad P(X_0 = S_2)]$$

$$\underline{\pi}_1 = [P(X_1 = S_0) \quad P(X_1 = S_1) \quad P(X_1 = S_2)]$$

$$\vdots$$

$$\underline{\pi}_n = [P(X_n = S_0) \quad P(X_n = S_1) \quad P(X_n = S_2)]$$

Q) How to find π_n from π_0?

Example: Two-State Markov Chain

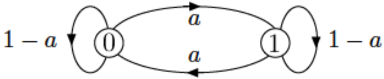
Here is a symmetric two-state Markov chain.

Example: Two-State Markov Chain

Here is a symmetric two-state Markov chain. It describes a random motion in $\{0, 1\}$.

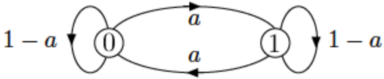
Example: Two-State Markov Chain

Here is a symmetric two-state Markov chain. It describes a random motion in $\{0, 1\}$. Here, *a* is the probability that the state changes in the next step.



Example: Two-State Markov Chain

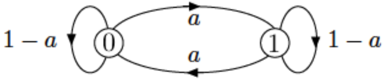
Here is a symmetric two-state Markov chain. It describes a random motion in $\{0, 1\}$. Here, *a* is the probability that the state changes in the next step.



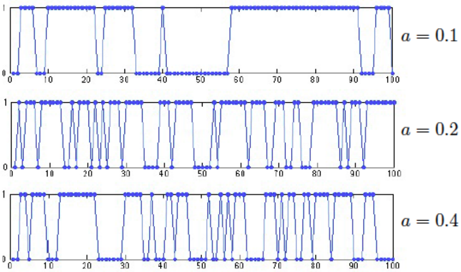
Let's simulate the Markov chain:

Example: Two-State Markov Chain

Here is a symmetric two-state Markov chain. It describes a random motion in $\{0, 1\}$. Here, *a* is the probability that the state changes in the next step.

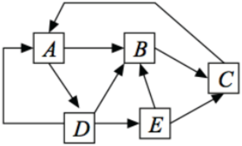


Let's simulate the Markov chain:



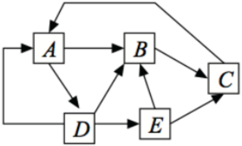
PageRank illustration: Five-State Markov Chain

At each step, the MC follows one of the outgoing arrows of the current state, with equal probabilities.



PageRank illustration: Five-State Markov Chain

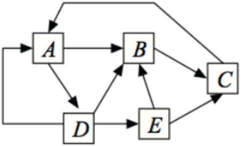
At each step, the MC follows one of the outgoing arrows of the current state, with equal probabilities.



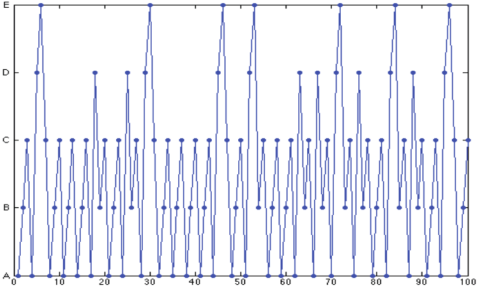
Let's simulate the Markov chain:

PageRank illustration: Five-State Markov Chain

At each step, the MC follows one of the outgoing arrows of the current state, with equal probabilities.

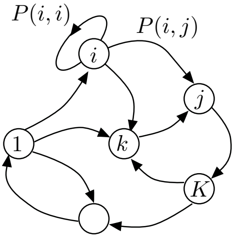


Let’s simulate the Markov chain:

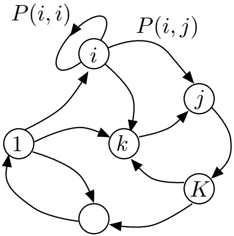


Finite Markov Chain: Definition

Finite Markov Chain: Definition

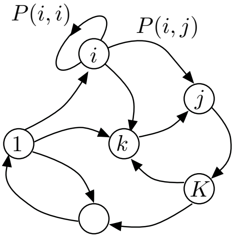


Finite Markov Chain: Definition



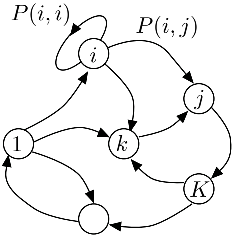
- A finite set of states: $\mathscr{X} = \{1, 2, \ldots, K\}$

Finite Markov Chain: Definition



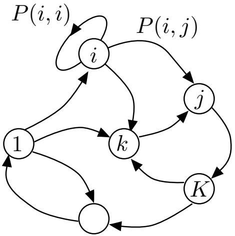
- A finite set of states: $\mathscr{X} = \{1, 2, \ldots, K\}$
- A probability distribution π_0 on $\mathscr{X}$:

Finite Markov Chain: Definition



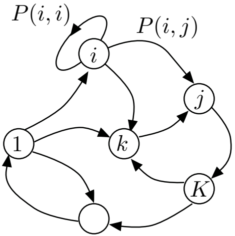
- A finite set of states: $\mathscr{X} = \{1, 2, \ldots, K\}$
- A probability distribution π_0 on $\mathscr{X}$: $\pi_0(i) \geq 0, \sum_i \pi_0(i) = 1$

Finite Markov Chain: Definition



- A finite set of states: $\mathscr{X} = \{1, 2, \ldots, K\}$
- A probability distribution π_0 on $\mathscr{X}$: $\pi_0(i) \geq 0, \sum_i \pi_0(i) = 1$
- Transition probabilities: $P(i,j)$ for $i,j \in \mathscr{X}$

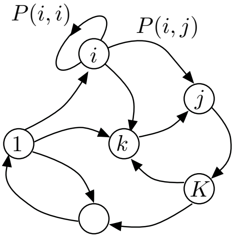
Finite Markov Chain: Definition



- A finite set of states: $\mathscr{X} = \{1, 2, \ldots, K\}$
- A probability distribution π_0 on $\mathscr{X}$: $\pi_0(i) \geq 0, \sum_i \pi_0(i) = 1$
- Transition probabilities: $P(i,j)$ for $i,j \in \mathscr{X}$

  $P(i,j) \geq 0, \forall i,j;$

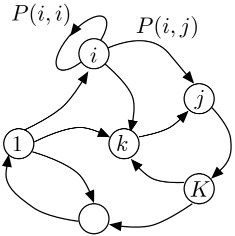
Finite Markov Chain: Definition



- A finite set of states: $\mathscr{X} = \{1, 2, \ldots, K\}$
- A probability distribution π_0 on $\mathscr{X}$: $\pi_0(i) \geq 0, \sum_i \pi_0(i) = 1$
- Transition probabilities: $P(i,j)$ for $i,j \in \mathscr{X}$

  $P(i,j) \geq 0, \forall i,j; \sum_j P(i,j) = 1, \forall i$

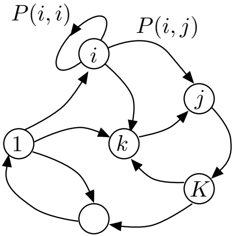
Finite Markov Chain: Definition



- A finite set of states: $\mathscr{X} = \{1, 2, \ldots, K\}$
- A probability distribution π_0 on $\mathscr{X}$: $\pi_0(i) \geq 0, \sum_i \pi_0(i) = 1$
- Transition probabilities: $P(i,j)$ for $i,j \in \mathscr{X}$

  $P(i,j) \geq 0, \forall i,j; \sum_j P(i,j) = 1, \forall i$
- $\{X_n, n \geq 0\}$ is defined so that

Finite Markov Chain: Definition



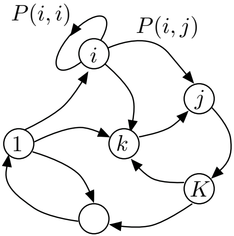
- A finite set of states: $\mathscr{X} = \{1, 2, \ldots, K\}$
- A probability distribution π_0 on $\mathscr{X}$: $\pi_0(i) \geq 0, \sum_i \pi_0(i) = 1$
- Transition probabilities: $P(i,j)$ for $i,j \in \mathscr{X}$

  $$P(i,j) \geq 0, \forall i,j; \sum_j P(i,j) = 1, \forall i$$

- $\{X_n, n \geq 0\}$ is defined so that

  $$Pr[X_0 = i] = \pi_0(i), i \in \mathscr{X}$$

Finite Markov Chain: Definition

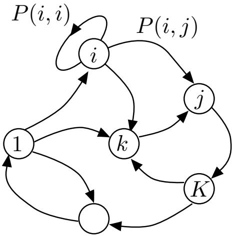


- A finite set of states: $\mathscr{X} = \{1, 2, \ldots, K\}$
- A probability distribution π_0 on $\mathscr{X}$: $\pi_0(i) \geq 0, \sum_i \pi_0(i) = 1$
- Transition probabilities: $P(i,j)$ for $i,j \in \mathscr{X}$

  $P(i,j) \geq 0, \forall i,j; \sum_j P(i,j) = 1, \forall i$
- $\{X_n, n \geq 0\}$ is defined so that

  $Pr[X_0 = i] = \pi_0(i), i \in \mathscr{X}$ (initial distribution)

Finite Markov Chain: Definition



- A finite set of states: $\mathscr{X} = \{1, 2, \ldots, K\}$
- A probability distribution π_0 on $\mathscr{X}$: $\pi_0(i) \geq 0, \sum_i \pi_0(i) = 1$
- Transition probabilities: $P(i,j)$ for $i,j \in \mathscr{X}$

  $P(i,j) \geq 0, \forall i,j; \sum_j P(i,j) = 1, \forall i$
- $\{X_n, n \geq 0\}$ is defined so that

  $Pr[X_0 = i] = \pi_0(i), i \in \mathscr{X}$ (initial distribution)

  $Pr[X_{n+1} = j \mid X_0, \ldots, X_n = i] = P(i,j), i,j \in \mathscr{X}.$

Irreducibility

Irreducibility

Definition A Markov chain is irreducible if it can go from every state *i* to every state *j*

Irreducibility

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

Irreducibility

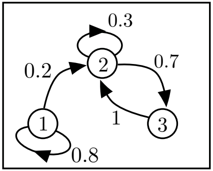
Definition A Markov chain is irreducible if it can go from every state i to every state j (possibly in multiple steps).

Examples:

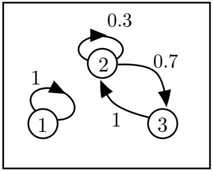
Irreducibility

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

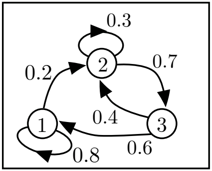
Examples:



[A]



[B]

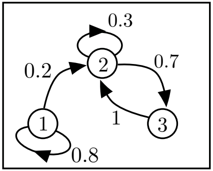


[C]

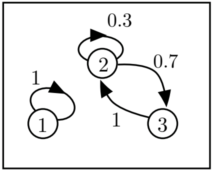
Irreducibility

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

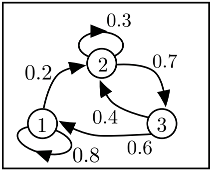
Examples:



[A]



[B]



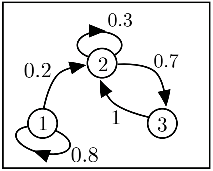
[C]

[A] is

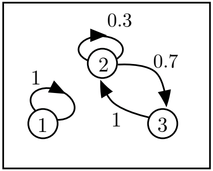
Irreducibility

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

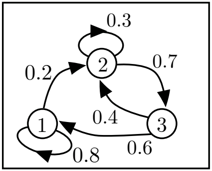
Examples:



[A]



[B]



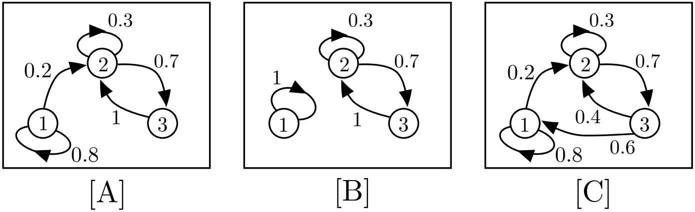
[C]

[A] is not irreducible.

Irreducibility

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

Examples:

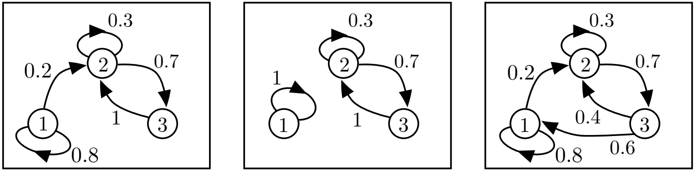


[A] is not irreducible. It cannot go from (2) to (1).

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]

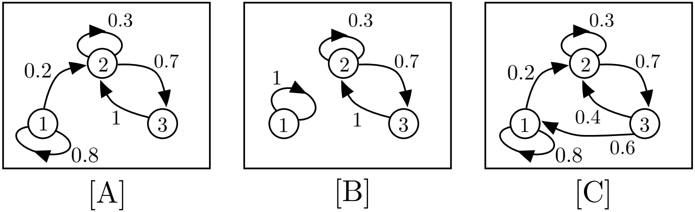
[A] is not irreducible. It cannot go from (2) to (1).

[B] is

Irreducibility

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

Examples:



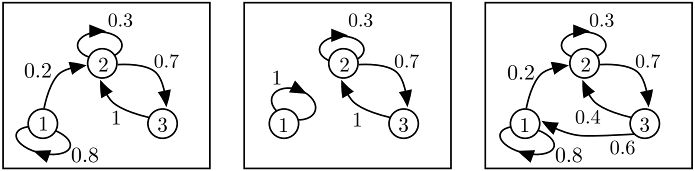
[A] is not irreducible. It cannot go from (2) to (1).

[B] is not irreducible.

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]

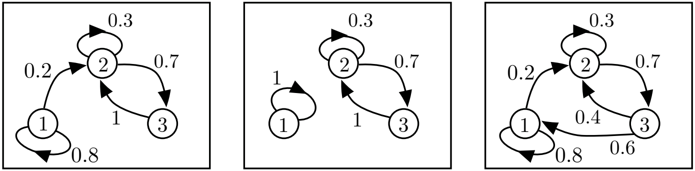
[A] is not irreducible. It cannot go from (2) to (1).

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

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).

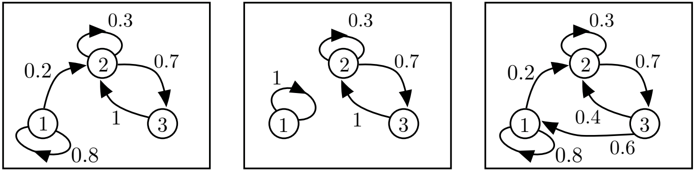
[B] is not irreducible. It cannot go from (2) to (1).

[C] is

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).

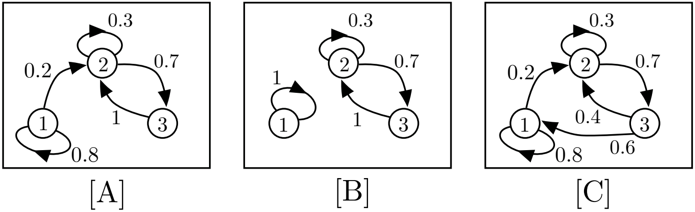
[B] is not irreducible. It cannot go from (2) to (1).

[C] is irreducible.

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

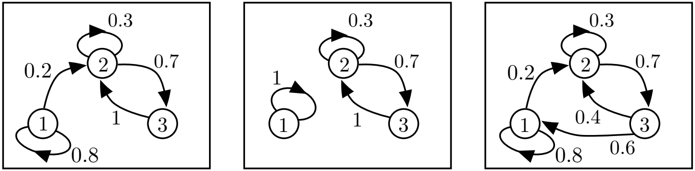
[B] is not irreducible. It cannot go from (2) to (1).

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

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).

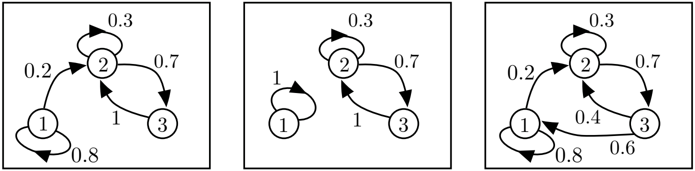
[C] is irreducible. It can go from every *i* to every *j*.

If you consider the graph with arrows when $P(i,j) > 0$,

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).

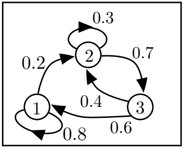
[B] is not irreducible. It cannot go from (2) to (1).

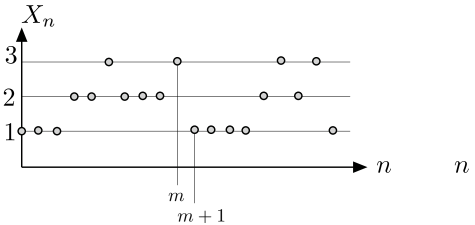
[C] is irreducible. It can go from every *i* to every *j*.

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

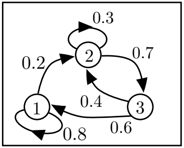
Finding π_n: the Distribution of X_n

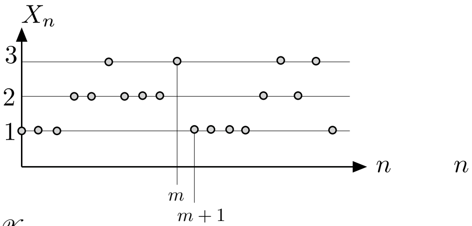
Finding π_n: the Distribution of X_n





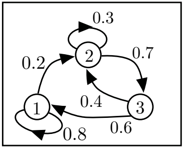
Finding π_n: the Distribution of X_n

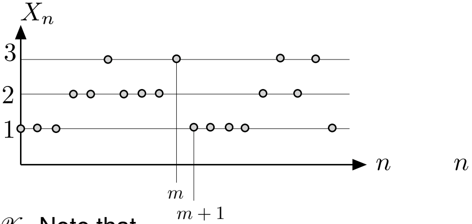




Let $\pi_m(i) = Pr[X_m = i], i \in \mathscr{X}$.

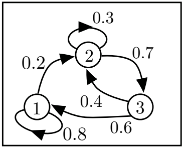
Finding π_n: the Distribution of X_n

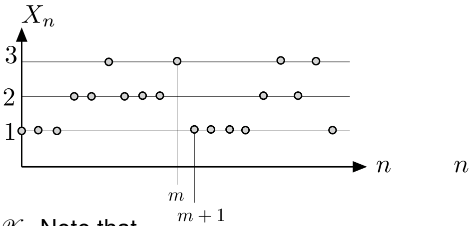




Let $\pi_m(i) = Pr[X_m = i], i \in \mathscr{X}$. Note that

Finding π_n: the Distribution of X_n

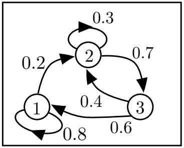


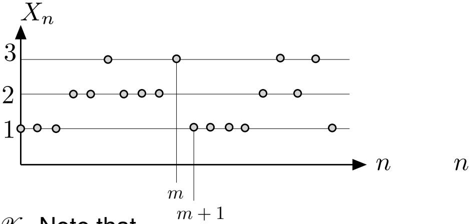


Let $\pi_m(i) = Pr[X_m = i], i \in \mathscr{X}$. Note that

$$Pr[X_{m+1} = j] = \sum_i Pr[X_{m+1} = j, X_m = i]$$

Finding π_n: the Distribution of X_n

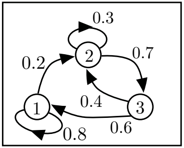


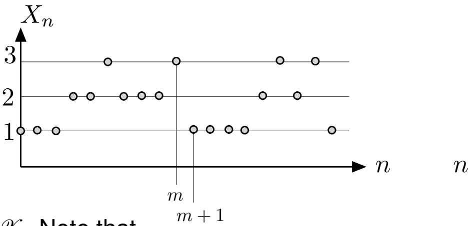


Let $\pi_m(i) = Pr[X_m = i], i \in \mathscr{X}$. Note that

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

Finding π_n: the Distribution of X_n

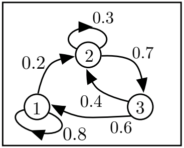


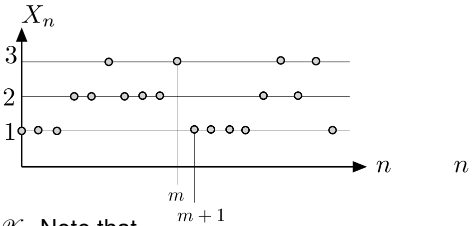


Let $\pi_m(i) = Pr[X_m = i], i \in \mathscr{X}$. Note that

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

Finding π_n: the Distribution of X_n





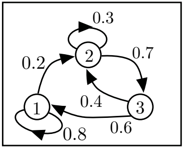
Let $\pi_m(i) = Pr[X_m = i], i \in \mathscr{X}$. Note that

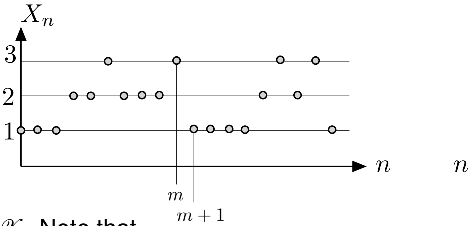
$$\begin{aligned} Pr[X_{m+1} = j] &= \sum_i Pr[X_{m+1} = j, X_m = i] \\ &= \sum_i Pr[X_m = i] Pr[X_{m+1} = j \mid X_m = i] \\ &= \sum_i \pi_m(i) P(i,j). \end{aligned}$$

Hence,

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

Finding π_n: the Distribution of X_n





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

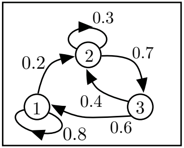
$$\begin{aligned} Pr[X_{m+1} = j] &= \sum_i Pr[X_{m+1} = j, X_m = i] \\ &= \sum_i Pr[X_m = i] Pr[X_{m+1} = j \mid X_m = i] \\ &= \sum_i \pi_m(i) P(i,j). \end{aligned}$$

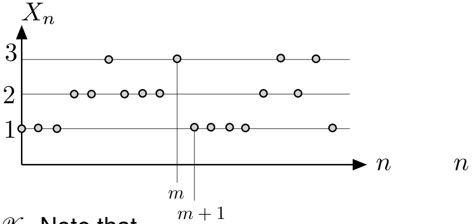
Hence,

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

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

Finding π_n: the Distribution of X_n





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

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

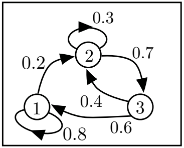
Hence,

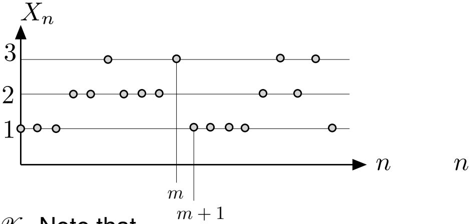
$$\pi_{m+1}(j) = \sum_i \pi_m(i) P(i,j), \forall j \in \mathscr{X}.$$

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

Thus, $\pi_1 = \pi_0 P$,

Finding π_n: the Distribution of X_n





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

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

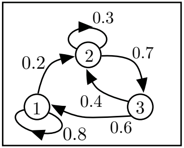
Hence,

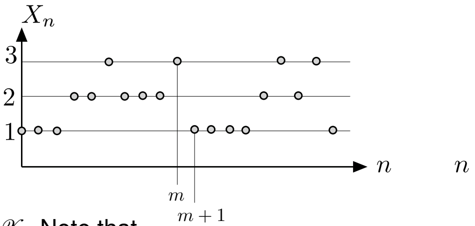
$$\pi_{m+1}(j) = \sum_i \pi_m(i) P(i,j), \forall j \in \mathscr{X}.$$

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

Thus, $\pi_1 = \pi_0 P$, $\pi_2 = \pi_1 P$

Finding π_n: the Distribution of X_n





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

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

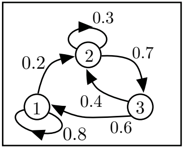
Hence,

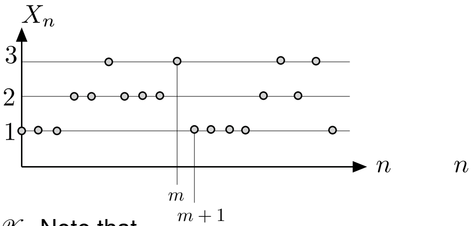
$$\pi_{m+1}(j) = \sum_i \pi_m(i) P(i,j), \forall j \in \mathscr{X}.$$

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

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

Finding π_n: the 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,

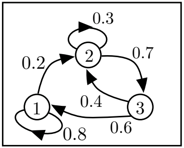
$$\pi_{m+1}(j) = \sum_i \pi_m(i) P(i,j), \forall j \in \mathscr{X}.$$

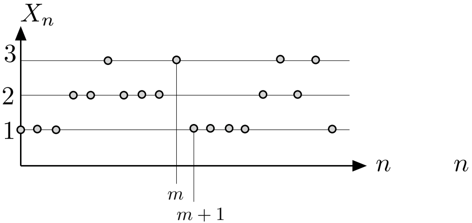
With π_m, π_{m+1} as row vectors, these identities are written as $\pi_{m+1} = \pi_m P$.

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

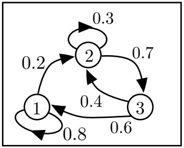
$$\pi_n = \pi_0 P^n, n \geq 0.$$

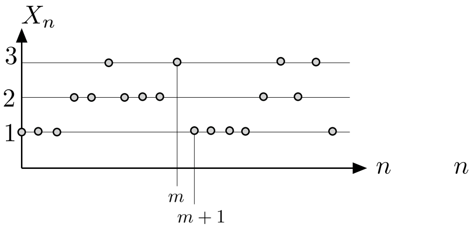
OH Ex.: Finding π_n, the distribution of X_n

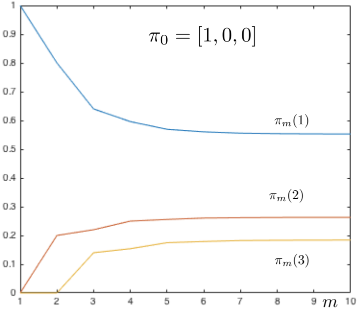


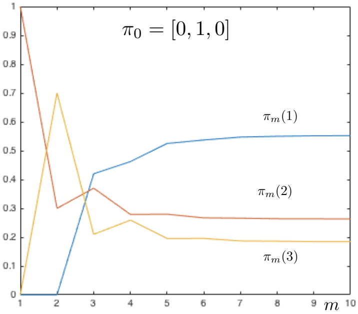


OH Ex.: Finding π_n, the distribution of X_n

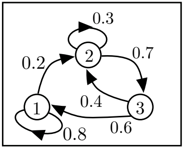


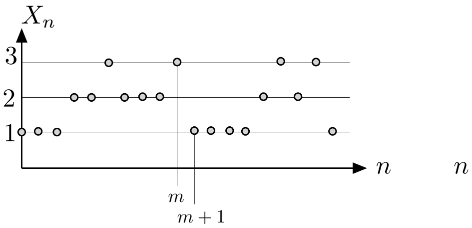


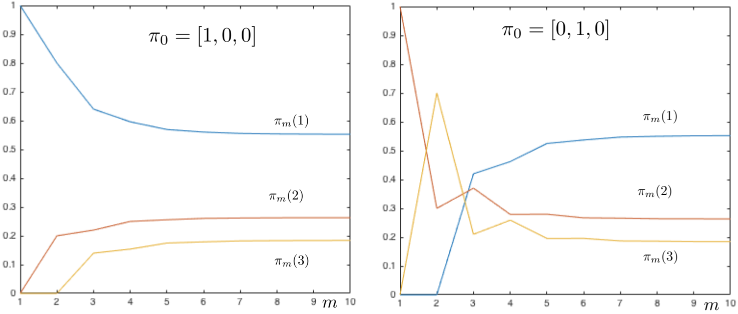




OH Ex.: Finding π_n, the distribution of X_n







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

Balance Equations

Balance Equations

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

Balance Equations

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

Defn.

Balance Equations

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

Defn. A distr. π_0 s.t. $\pi_m = \pi_0, \forall m$ is called an invariant distribution.

Balance Equations

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

Defn. A distr. π_0 s.t. $\pi_m = \pi_0, \forall m$ is called an invariant distribution.

Theorem

Balance Equations

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

Defn. A distr. π_0 s.t. $\pi_m = \pi_0, \forall m$ is called an invariant distribution.

Theorem A distribution π_0 is invariant iff $\pi_0 P = \pi_0$.

Balance Equations

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

Defn. A distr. π_0 s.t. $\pi_m = \pi_0, \forall m$ is called an invariant distribution.

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

Balance Equations

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

Defn. A distr. π_0 s.t. $\pi_m = \pi_0, \forall m$ is called an invariant distribution.

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

Balance Equations

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

Defn. A distr. π_0 s.t. $\pi_m = \pi_0, \forall m$ is called an invariant distribution.

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

If π_0 is invariant, the distr. of X_n is the same as that of X_0.

Balance Equations

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

Defn. A distr. π_0 s.t. $\pi_m = \pi_0, \forall m$ is called an invariant distribution.

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

If π_0 is invariant, the distr. of X_n is the same as that of X_0. Of course, this does **not** mean that nothing moves.

Balance Equations

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

Defn. A distr. π_0 s.t. $\pi_m = \pi_0, \forall m$ is called an invariant distribution.

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

If π_0 is invariant, the distr. of X_n is the same as that of X_0. Of course, this does **not** mean that nothing moves. It means that **prob. flow** leaving state i = **prob. flow** entering state i; $\forall i \in \mathscr{X}$.

Balance Equations

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

Defn. A distr. π_0 s.t. $\pi_m = \pi_0, \forall m$ is called an invariant distribution.

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

If π_0 is invariant, the distr. of X_n is the same as that of X_0. Of course, this does **not** mean that nothing moves. It means that **prob. flow** leaving state i = **prob. flow** entering state i; $\forall i \in \mathscr{X}$. That is, **Prob. flow out = Prob. flow in for all states in the MC.**

Recall, the state transition equations from earlier slide:

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

Balance Equations

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

Defn. A distr. π_0 s.t. $\pi_m = \pi_0, \forall m$ is called an invariant distribution.

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

If π_0 is invariant, the distr. of X_n is the same as that of X_0. Of course, this does **not** mean that nothing moves. It means that **prob. flow** leaving state i = **prob. flow** entering state i; $\forall i \in \mathscr{X}$. That is,
Prob. flow out = Prob. flow in for all states in the MC.

Recall, the state transition equations from earlier slide:

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

The balance equations say that

Balance Equations

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

Defn. A distr. π_0 s.t. $\pi_m = \pi_0, \forall m$ is called an invariant distribution.

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

If π_0 is invariant, the distr. of X_n is the same as that of X_0. Of course, this does **not** mean that nothing moves. It means that **prob. flow** leaving state i = **prob. flow** entering state i; $\forall i \in \mathscr{X}$. That is,
Prob. flow out = Prob. flow in for all states in the MC.

Recall, the state transition equations from earlier slide:

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

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

Balance Equations

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

Defn. A distr. π_0 s.t. $\pi_m = \pi_0, \forall m$ is called an invariant distribution.

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

If π_0 is invariant, the distr. of X_n is the same as that of X_0. Of course, this does **not** mean that nothing moves. It means that **prob. flow** leaving state i = **prob. flow** entering state i; $\forall i \in \mathscr{X}$. That is,
Prob. flow out = Prob. flow in for all states in the MC.

Recall, the state transition equations from earlier slide:

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

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

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

Balance Equations

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

Defn. A distr. π_0 s.t. $\pi_m = \pi_0, \forall m$ is called an invariant distribution.

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

If π_0 is invariant, the distr. of X_n is the same as that of X_0. Of course, this does **not** mean that nothing moves. It means that **prob. flow** leaving state i = **prob. flow** entering state i; $\forall i \in \mathscr{X}$. That is,
Prob. flow out = Prob. flow in for all states in the MC.

Recall, the state transition equations from earlier slide:

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

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

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

Balance Equations

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

Defn. A distr. π_0 s.t. $\pi_m = \pi_0, \forall m$ is called an invariant distribution.

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

If π_0 is invariant, the distr. of X_n is the same as that of X_0. Of course, this does **not** mean that nothing moves. It means that **prob. flow** leaving state i = **prob. flow** entering state i; $\forall i \in \mathscr{X}$. That is,
Prob. flow out = Prob. flow in for all states in the MC.

Recall, the state transition equations from earlier slide:

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

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

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

Thus, (LHS=) $Pr[\text{enter } i] = (RHS =) Pr[\text{leave } i]$.

Invariant Distribution: always exist?

Question 1:

Invariant Distribution: always exist?

Question 1: Does a MC *always* have an invariant distribution?

Invariant Distribution: always exist?

Question 1: Does a MC *always* have an invariant distribution?

Question 2:

Invariant Distribution: always exist?

Question 1: Does a MC *always* have an invariant distribution?

Question 2: If an invariant distribution exists, is it unique?

Invariant Distribution: always exist?

Question 1: Does a MC *always* have an invariant distribution?

Question 2: If an invariant distribution exists, is it unique?

Answer 1: If the number of states in the MC is **finite**, then the answer to Question 1 is **yes.**

Invariant Distribution: always exist?

Question 1: Does a MC *always* have an invariant distribution?

Question 2: If an invariant distribution exists, is it unique?

Answer 1: If the number of states in the MC is **finite**, then the answer to Question 1 is **yes.**

Answer 2: If the MC is finite and **irreducible**, then the answer to Question 2 is **yes.**

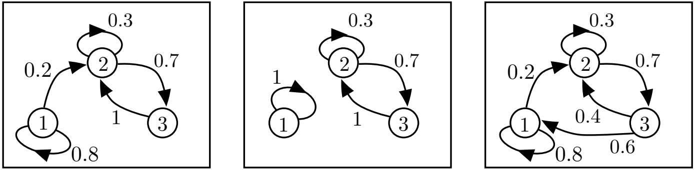
Invariant Distribution: always exist?

Question 1: Does a MC *always* have an invariant distribution?

Question 2: If an invariant distribution exists, is it unique?

Answer 1: If the number of states in the MC is **finite**, then the answer to Question 1 is **yes.**

Answer 2: If the MC is finite and **irreducible**, then the answer to Question 2 is **yes.**



[A] [B] [C]

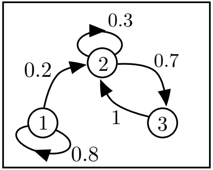
Invariant Distribution: always exist?

Question 1: Does a MC *always* have an invariant distribution?

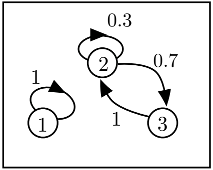
Question 2: If an invariant distribution exists, is it unique?

Answer 1: If the number of states in the MC is **finite**, then the answer to Question 1 is **yes.**

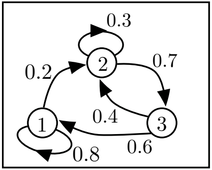
Answer 2: If the MC is finite and **irreducible**, then the answer to Question 2 is **yes.**



[A]



[B]



[C]

Proof:

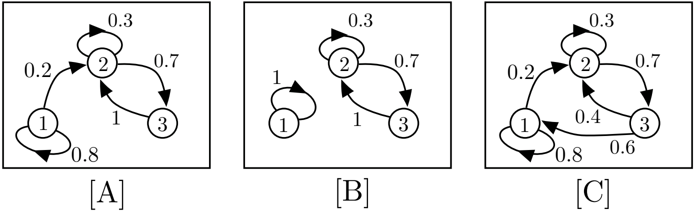
Invariant Distribution: always exist?

Question 1: Does a MC *always* have an invariant distribution?

Question 2: If an invariant distribution exists, is it unique?

Answer 1: If the number of states in the MC is **finite**, then the answer to Question 1 is **yes.**

Answer 2: If the MC is finite and **irreducible**, then the answer to Question 2 is **yes.**



Proof: (EECS 126)

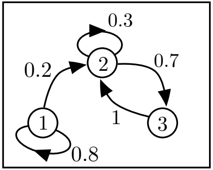
Invariant Distribution: always exist?

Question 1: Does a MC *always* have an invariant distribution?

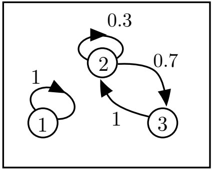
Question 2: If an invariant distribution exists, is it unique?

Answer 1: If the number of states in the MC is **finite**, then the answer to Question 1 is **yes.**

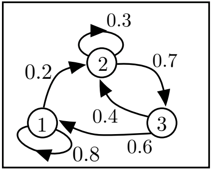
Answer 2: If the MC is finite and **irreducible**, then the answer to Question 2 is **yes.**



[A]



[B]



[C]

Proof: (EECS 126)

Other settings? (e.g. infinite chains, periodicity,...?)

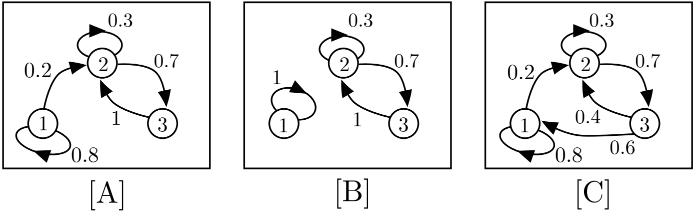
Invariant Distribution: always exist?

Question 1: Does a MC *always* have an invariant distribution?

Question 2: If an invariant distribution exists, is it unique?

Answer 1: If the number of states in the MC is **finite**, then the answer to Question 1 is **yes.**

Answer 2: If the MC is finite and **irreducible**, then the answer to Question 2 is **yes.**



Proof: (EECS 126)

Other settings? (e.g. infinite chains, periodicity,...?) (EECS 126)

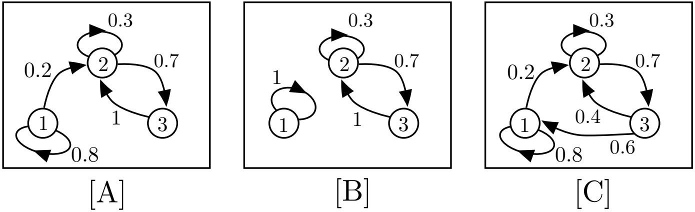
Invariant Distribution: always exist?

Question 1: Does a MC *always* have an invariant distribution?

Question 2: If an invariant distribution exists, is it unique?

Answer 1: If the number of states in the MC is **finite**, then the answer to Question 1 is **yes.**

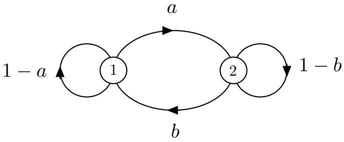
Answer 2: If the MC is finite and **irreducible**, then the answer to Question 2 is **yes.**



Proof: (EECS 126)

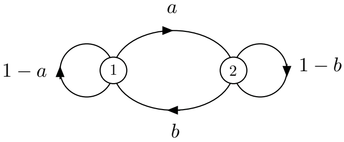
Other settings? (e.g. infinite chains, periodicity,...?) (EECS 126)

Balance Equations: 2-state MC example



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

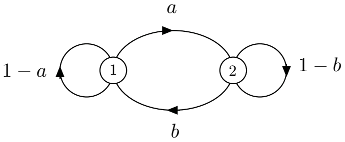
Balance Equations: 2-state MC example



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

$$\pi P = \pi$$

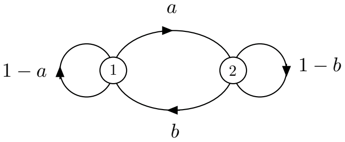
Balance Equations: 2-state MC example



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

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

Balance Equations: 2-state MC example

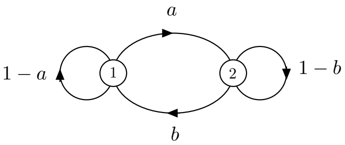


$$P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}$$

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

$$\Leftrightarrow \quad \pi(1)(1-a) + \pi(2)b = \pi(1) \text{ and}$$

Balance Equations: 2-state MC example

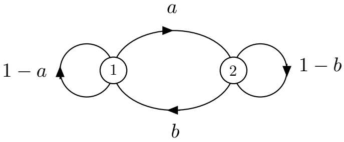


$$P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}$$

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

$$\Leftrightarrow \quad \pi(1)(1-a) + \pi(2)b = \pi(1) \text{ and } \pi(1)a + \pi(2)(1-b) = \pi(2)$$

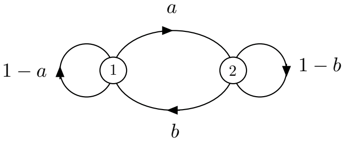
Balance Equations: 2-state MC example



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

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

Balance Equations: 2-state MC example

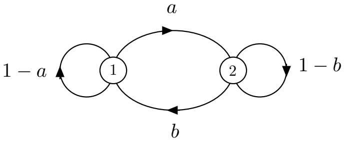


$$P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}$$

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

Prob. flow leaving state 1 = Prob. flow entering state 1

Balance Equations: 2-state MC example



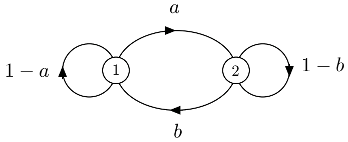
$$P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}$$

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

Prob. flow leaving state 1 = Prob. flow entering state 1

These equations are redundant!

Balance Equations: 2-state MC example



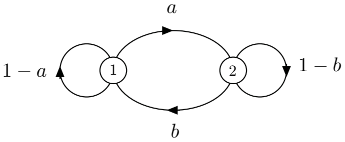
$$P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}$$

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

Prob. flow leaving state 1 = Prob. flow entering state 1

These equations are redundant! We have to add an equation:

Balance Equations: 2-state MC example



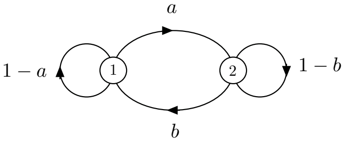
$$P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}$$

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

Prob. flow leaving state 1 = Prob. flow entering state 1

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

Balance Equations: 2-state MC example



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

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

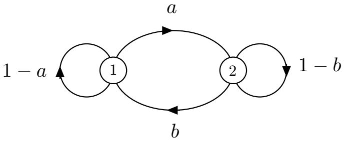
$$\Leftrightarrow \quad \pi(1)(1-a) + \pi(2)b = \pi(1) \text{ and } \pi(1)a + \pi(2)(1-b) = \pi(2)$$

$$\Leftrightarrow \quad \pi(1)a = \pi(2)b.$$

Prob. flow leaving state 1 = Prob. flow entering state 1

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

Balance Equations: 2-state MC example



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

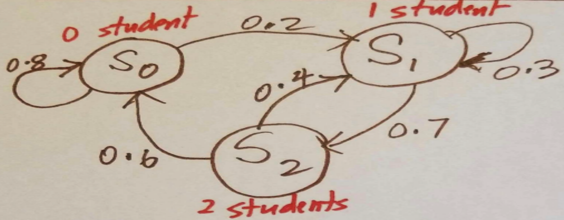
$$\begin{aligned} \pi P = \pi \quad & \Leftrightarrow \quad [\pi(1), \pi(2)] \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} = [\pi(1), \pi(2)] \\ & \Leftrightarrow \quad \pi(1)(1-a) + \pi(2)b = \pi(1) \text{ and } \pi(1)a + \pi(2)(1-b) = \pi(2) \\ & \Leftrightarrow \quad \pi(1)a = \pi(2)b. \end{aligned}$$

Prob. flow leaving state 1 = Prob. flow entering state 1

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

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

OH Example: Balance Eqs.



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

$$\boxed{\underline{\pi}_0 P = \underline{\pi}_0}$$ (Balance Eqs.)

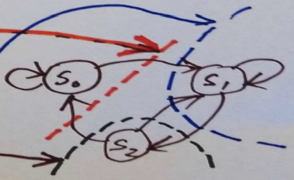
Let $\underline{\pi}_0 = [x \quad y \quad z]$

"FLOW IN = FLOW OUT":

S_0: $0.6z = 0.2x$

S_1: $0.2x + 0.4z = 0.7y$

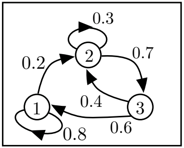
S_2: $0.7y = z$

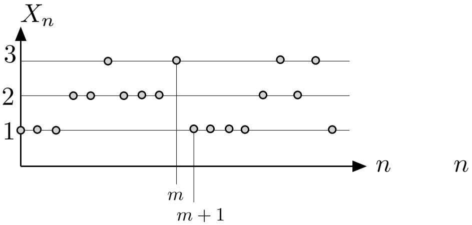


Need one more equation as one eqn. above is redundant:

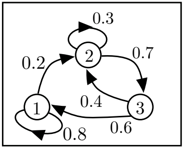
$\boxed{x + y + z = 1}$ ⇒ Solve for x, y, z ⇒ $\boxed{\underline{\pi}_0 \cong [0.55 \quad 0.26 \quad 0.19]}$

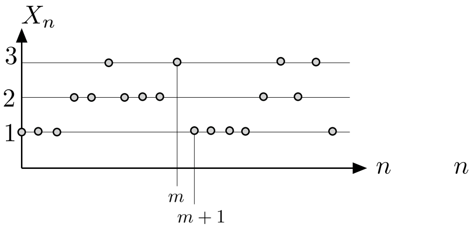
Finding π_n: the Distribution of X_n

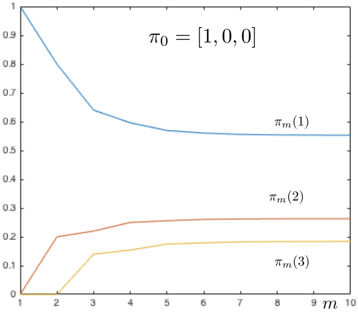


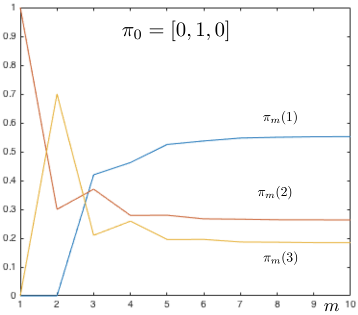


Finding π_n: the Distribution of X_n

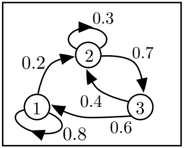


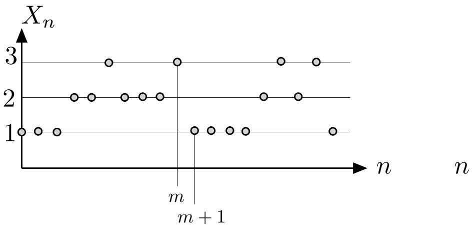


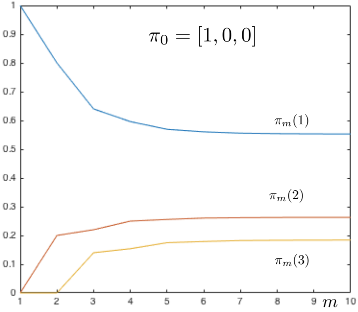


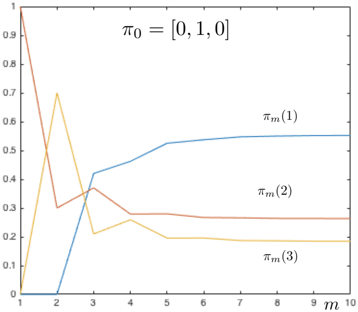


Finding π_n: the Distribution of X_n









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

Summary

Markov Chains

Summary

Markov Chains

Summary

Markov Chains

1. Random Process: **sequence** of Random Variables;

Summary

Markov Chains

1. Random Process: **sequence** of Random Variables;
2. Markov Chain: $Pr[X_{n+1} = j \mid X_0, \ldots, X_n = i] = P(i,j), i,j \in \mathscr{X}$

Summary

Markov Chains

1. Random Process: **sequence** of Random Variables;
2. Markov Chain: $Pr[X_{n+1} = j \mid X_0, \ldots, X_n = i] = P(i,j), i,j \in \mathscr{X}$
3. Invariant Distribution of Markov Chain: balance equations

Summary

Markov Chains

1. Random Process: **sequence** of Random Variables;
2. Markov Chain: $Pr[X_{n+1} = j \mid X_0, \ldots, X_n = i] = P(i,j), i,j \in \mathscr{X}$
3. Invariant Distribution of Markov Chain: balance equations