Markov chains and applications

Dr. András Pongrácz

Week 5, University of Debrecen

Finite Markov chains

... as digraphs

S is a finite set, $|S| = n$

... as digraphs

S is a finite set, $|S| = n$ (states)

S is a finite set, $|S| = n$ (states)

A weighted digraph with set of vertices S is a Markov chain if all weights are non-negative and the sum of weights on all outgoing edges from any given vertex is 1.

... as digraphs

S is a finite set, $|S| = n$ (states)

A weighted digraph with set of vertices S is a Markov chain if all weights are non-negative and the sum of weights on all outgoing edges from any given vertex is 1.

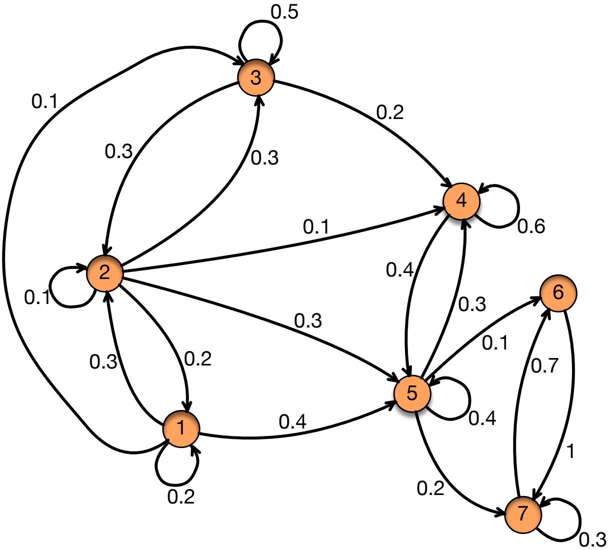
The weights are the transition probabilities.

S is a finite set, $|S| = n$ (states)

A weighted digraph with set of vertices S is a Markov chain if all weights are non-negative and the sum of weights on all outgoing edges from any given vertex is 1.

The weights are the transition probabilities.

The idea is that we walk on the states, and at every step we decide where to move according to the probability distribution on the outgoing edges.



0.5
3
0.1
0.2
0.3
0.3
4
0.6
0.1
2
0.4
0.1
0.3
6
0.3
0.1
0.3
0.2
0.7
5
0.4
1
0.4
1
0.2
0.2
7
0.3

Too complicated, intransparent.

Too complicated, intransparent.

Better representation: transition matrix P.

Too complicated, intransparent.

Better representation: transition matrix P.

$P[i,j]$ is the transition probability from s_i to s_j.

Too complicated, intransparent.

Better representation : transition matrix P.

$P[i, j]$ is the transition probability from s_i to s_j. I.e., it is the probability of moving to s_j in one step, provided we are in s_i.

Too complicated, intransparent.

Better representation: transition matrix P.

$P[i, j]$ is the transition probability from s_i to s_j. I.e., it is the probability of moving to s_j in one step, provided we are in s_i. Note that this really depends only on i, j (it is irrelevant how we arrived at s_i).

Too complicated, intransparent.

Better representation: transition matrix P.

$P[i,j]$ is the transition probability from s_i to s_j. I.e., it is the probability of moving to s_j in one step, provided we are in s_i. Note that this really depends only on i,j (it is irrelevant how we arrived at s_i).
Basic properties:

$$P[i,j] \geq 0$$

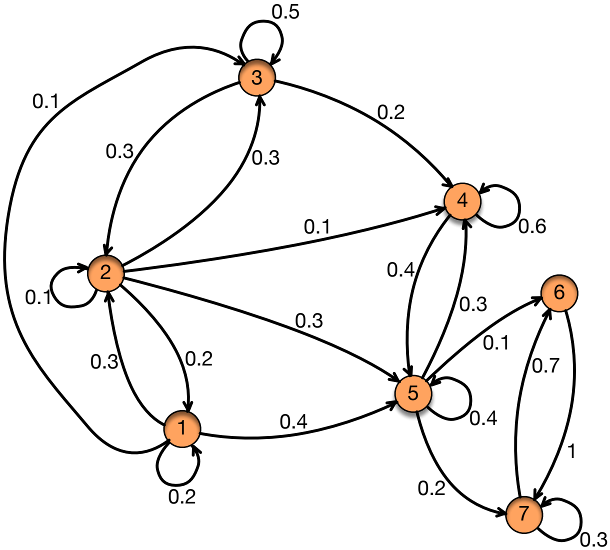
$$\sum_{j=1}^{n} P[i,j] = 1 \text{ for all } 1 \leq i \leq n.$$

Too complicated, intransparent.

Better representation : transition matrix P.

$P[i,j]$ is the transition probability from s_i to s_j. I.e., it is the probability of moving to s_j in one step, provided we are in s_i. Note that this really depends only on i,j (it is irrelevant how we arrived at s_i).
Basic properties :

- $P[i,j] \geq 0$
- $\sum_{j=1}^{n} P[i,j] = 1$ for all $1 \leq i \leq n$. Equivalently $P\underline{1} = \underline{1}$, so the all-1 vector is an eigenvector of P with eigenvalue 1.


0.5
3
0.1
0.2
0.3
0.3
4
0.6
0.1
2
0.4
0.1
0.3
6
0.3
0.1
0.3
0.2
0.7
5
0.4
1
0.4
1
0.2
0.2
7
0.3

$$P = \begin{pmatrix} 0.2 & 0.3 & 0.1 & 0 & 0.4 & 0 & 0 \\ 0.2 & 0.1 & 0.3 & 0.1 & 0.3 & 0 & 0 \\ 0 & 0.3 & 0.5 & 0.2 & 0 & 0 & 0 \\ 0 & 0 & 0.6 & 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0 & 0.3 & 0.4 & 0.1 & 0.2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0.7 & 0.3 \end{pmatrix}$$

Powers of the transition matrix

How do we get from s_i to s_j in 2 steps?

Powers of the transition matrix

How do we get from s_i to s_j in 2 steps? Via a third state s_k.

Powers of the transition matrix

How do we get from s_i to s_j in 2 steps? Via a third state s_k. So the probability is

$$\sum_{k=1}^{n} P[i,k]P[k,j] = (P^2)[i,j]$$

Powers of the transition matrix

How do we get from s_i to s_j in 2 steps? Via a third state s_k. So the probability is

$$\sum_{k=1}^{n} P[i,k]P[k,j] = (P^2)[i,j]$$

A similar formula holds in general.

Powers of the transition matrix

How do we get from s_i to s_j in 2 steps? Via a third state s_k. So the probability is

$$\sum_{k=1}^{n} P[i,k]P[k,j] = (P^2)[i,j]$$

A similar formula holds in general.

Proposition

The probability of being at s_j after m steps starting from s_i is

$$(P^m)[i,j]$$

Induced digraph on strong components

Given the strong components $C_1, \ldots, C_k$ of a digraph, we define a digraph on them:

Induced digraph on strong components

Given the strong components $C_1, \ldots, C_k$ of a digraph, we define a digraph on them:
$\exists$ edge between C_i and C_j iff there is at least one such edge in the original digraph.

Induced digraph on strong components

Given the strong components $C_1, \ldots, C_k$ of a digraph, we define a digraph on them:
$\exists$ edge between C_i and C_j iff there is at least one such edge in the original digraph.

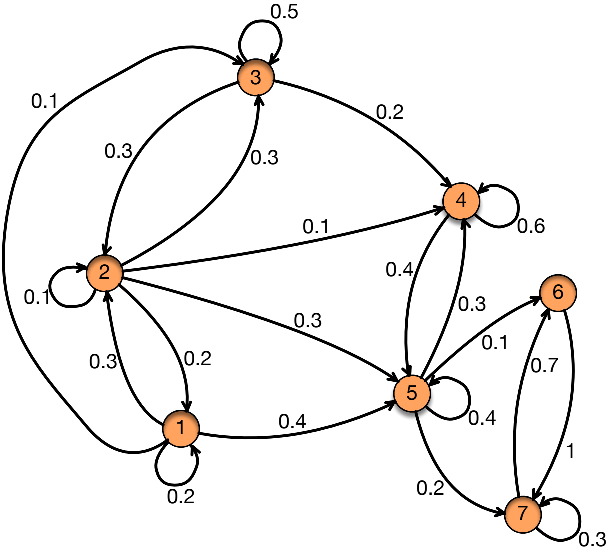
This yields an acyclic digraph (no directed cycle).

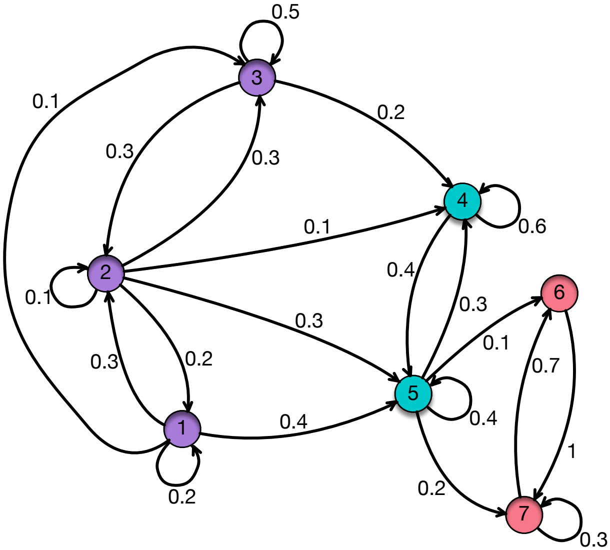
Induced digraph on strong components

Given the strong components $C_1, \ldots, C_k$ of a digraph, we define a digraph on them:
$\exists$ edge between C_i and C_j iff there is at least one such edge in the original digraph.

This yields an acyclic digraph (no directed cycle).

We can deconstruct the walk into two walks: first on this acyclic graph, until we reach a sink, and then a walk on the sink. (Cf. topological sorting of an acyclic digraph.)


0.5
3
0.1
0.2
0.3
0.3
4
0.6
0.1
0.4
2
0.1
0.3
6
0.3
0.1
0.3
0.2
0.7
5
0.4
1
0.4
1
0.2
0.2
7
0.3


0.5
3
0.1
0.3
0.3
0.2
4
0.6
0.1
2
0.1
0.4
0.3
6
0.3
0.3
0.1
0.3
0.2
0.7
5
0.4
1
0.4
1
0.2
0.2
7
0.3

Collapsing the sinks to one vertex leads to a very special type of chains.

Absorbing chains

Collapsing the sinks to one vertex leads to a very special type of chains.

Absorption

A state in a Markov chain is absorbing if it has only one outgoing edge, and the edge points to the state (and then the transition probability assigned to it is necessarily 1).

Collapsing the sinks to one vertex leads to a very special type of chains.

Absorption

A state in a Markov chain is absorbing if it has only one outgoing edge, and the edge points to the state (and then the transition probability assigned to it is necessarily 1). A finite Markov chain is absorbing if (it has at least one absorbing state and) from any state, we can reach an absorbing state with positive probability in a finite number of steps.

So every finite Markov chain is the composition of an absorbing chain and an irreducible one.

So every finite Markov chain is the composition of an absorbing chain and an irreducible one.

These two notions are very much on opposite ends of a spectrum.

So every finite Markov chain is the composition of an absorbing chain and an irreducible one.

These two notions are very much on opposite ends of a spectrum.

The natural problems to study on them are completely different,

So every finite Markov chain is the composition of an absorbing chain and an irreducible one.

These two notions are very much on opposite ends of a spectrum.

The natural problems to study on them are completely different, so they are investigated separately on the next classes.

Exercises

1. Show that if P is the transition matrix of a Markov chain, then all Gershgorin discs of P have a common point.

Exercises

1. Show that if P is the transition matrix of a Markov chain, then all Gershgorin discs of P have a common point.
2. Prove the proposition about $P^m[i, j]$.

Exercises

1. Show that if P is the transition matrix of a Markov chain, then all Gershgorin discs of P have a common point.
2. Prove the proposition about $P^m[i,j]$.
3. Show that a Markov chain is absorbing iff all sinks of the induced digraph on its strong components are singletons.

Exercises

1. Show that if P is the transition matrix of a Markov chain, then all Gershgorin discs of P have a common point.
2. Prove the proposition about $P^m[i,j]$.
3. Show that a Markov chain is absorbing iff all sinks of the induced digraph on its strong components are singletons.
4. Show that a matrix P is a transition matrix (of some Markov chain) iff P is non-negative and $P\underline{1} = \underline{1}$.

Exercises

1. Show that if P is the transition matrix of a Markov chain, then all Gershgorin discs of P have a common point.
2. Prove the proposition about $P^m[i,j]$.
3. Show that a Markov chain is absorbing iff all sinks of the induced digraph on its strong components are singletons.
4. Show that a matrix P is a transition matrix (of some Markov chain) iff P is non-negative and $P\underline{1} = \underline{1}$.
5. What is the difference between acyclic graphs and partial orders?

Exercises

1. Show that if P is the transition matrix of a Markov chain, then all Gershgorin discs of P have a common point.
2. Prove the proposition about $P^m[i,j]$.
3. Show that a Markov chain is absorbing iff all sinks of the induced digraph on its strong components are singletons.
4. Show that a matrix P is a transition matrix (of some Markov chain) iff P is non-negative and $P\underline{1} = \underline{1}$.
5. What is the difference between acyclic graphs and partial orders?
6. Prove that if P is absorbing, then so is P^m for any m. What does an absorbing, primitive chain look like?