Markov chains and applications

Dr. András Pongrácz

Week 10, University of Debrecen

Reversible chains

Definition

Given an irreducible Markov chain with transition matrix P and stationary distribution $\underline{w}^* = (w_1, \ldots, w_n)$.

Definition

Given an irreducible Markov chain with transition matrix P and stationary distribution $\underline{w}^* = (w_1, \ldots, w_n)$. The reversed chain is defined by the transition matrix $\widehat{P}$ with transition probabilities

$$\widehat{P}[i,k] := \frac{w_k}{w_i} P[k,i]$$

.

Definition

Given an irreducible Markov chain with transition matrix P and stationary distribution $\underline{w}^* = (w_1, \ldots, w_n)$. The reversed chain is defined by the transition matrix $\widehat{P}$ with transition probabilities

$$\widehat{P}[i,k] := \frac{w_k}{w_i} P[k,i]$$

.

Then $\widehat{P}$ is also irreducible, and $\widehat{\widehat{P}} = P$; cf. the exercises.

In what sense does $\widehat{P}$ reverse P?

In what sense does $\widehat{P}$ reverse P?
An N-step walk on the states of P is a process $(X_1, \ldots, X_N)$.

In what sense does $\widehat{P}$ reverse P?

An N-step walk on the states of P is a process $(X_1, \ldots, X_N)$. Assume that X_1 is sampled from the stationary distribution $\underline{w}^*$;

In what sense does $\widehat{P}$ reverse P?

An N-step walk on the states of P is a process $(X_1, \ldots, X_N)$. Assume that X_1 is sampled from the stationary distribution $\underline{w}^*$; then so are all the X_j.

In what sense does $\widehat{P}$ reverse P?

An N-step walk on the states of P is a process $(X_1, \ldots, X_N)$. Assume that X_1 is sampled from the stationary distribution $\underline{w}^*$; then so are all the X_j. What is the probability that we started from the state s_k under the condition that we ended up at state s_i in one move?

In what sense does $\widehat{P}$ reverse P?

An N-step walk on the states of P is a process $(X_1, \ldots, X_N)$. Assume that X_1 is sampled from the stationary distribution $\underline{w}^*$; then so are all the X_j. What is the probability that we started from the state s_k under the condition that we ended up at state s_i in one move? In one step:

$$\mathbb{P}(X_j = s_k \mid X_{j+1} = s_i) = \frac{\mathbb{P}(X_j = s_k \wedge X_{j+1} = s_i)}{\mathbb{P}(X_{j+1} = s_i)} = \frac{w_k P[k, i]}{w_i} = \widehat{P}[i, k]$$

In what sense does $\widehat{P}$ reverse P?
An N-step walk on the states of P is a process $(X_1, \dots, X_N)$. Assume that X_1 is sampled from the stationary distribution $\underline{w}^*$; then so are all the X_j. What is the probability that we started from the state s_k under the condition that we ended up at state s_i in one move? In one step:

$$\mathbb{P}(X_j = s_k \mid X_{j+1} = s_i) = \frac{\mathbb{P}(X_j = s_k \wedge X_{j+1} = s_i)}{\mathbb{P}(X_{j+1} = s_i)} = \frac{w_k P[k, i]}{w_i} = \widehat{P}[i, k]$$

Clearly $\sum_{k=1}^{n} \widehat{P}[i, k] = \sum_{k=1}^{n} \mathbb{P}(X_1 = s_k \mid X_2 = s_i) = 1$, and all the $\widehat{P}[i, k]$ are non-negative, thus $\widehat{P}$ is the transition matrix of some Markov chain.

Reversibility and balance

Definition

An irreducible Markov chain with transition matrix P is reversible if $P = \widehat{P}$.

Reversibility and balance

Definition

An irreducible Markov chain with transition matrix P is reversible if $P = \widehat{P}$.

Definition

An irreducible Markov chain with transition matrix P on n states satisfies the detailed balance equations for a nonzero row vector $\underline{\mu}^*$ if for all pairs $1 \leq i, k \leq n$ we have $\mu_i P[i,k] = \mu_k P[k,i]$.

Reversibility and balance

Definition

An irreducible Markov chain with transition matrix P is reversible if $P = \widehat{P}$.

Definition

An irreducible Markov chain with transition matrix P on n states satisfies the detailed balance equations for a nonzero row vector $\underline{\mu}^*$ if for all pairs $1 \leq i, k \leq n$ we have $\mu_i P[i,k] = \mu_k P[k,i]$. More generally, P satisfies the detailed balance equation for $\underline{\mu}^*$ and $n \times n$ square matrix Q if $\mu_i Q[i,k] = \mu_k P[k,i]$.

Reversibility and balance

Clearly, if P is reversible, then the detailed balance equations hold for the stationary distribution $\underline{\mu}^* = \underline{\boldsymbol{w}}^*$ and $Q = \widehat{P}$.

Clearly, if P is reversible, then the detailed balance equations hold for the stationary distribution $\underline{\mu}^* = \underline{w}^*$ and $Q = \widehat{P}$. The converse is Kelly's lemma.

Kelly's lemma

Let $P, Q \in M_n(\mathbb{R})$ be stochastic matrices and let $\underline{\mu}^*$ be a probability distribution such that the detailed balance equations hold. Then $\underline{\mu}^* = \underline{w}^*$ is the stationary distribution of P and $Q = \widehat{P}$.

Proof

For any i we have

$$\sum_{k=1}^{n} \mu_k P[k,i] = \sum_{k=1}^{n} \mu_i Q[i,k] = \mu_i \sum_{k=1}^{n} Q[i,k] = \mu_i.$$

Proof

For any i we have

$$\sum_{k=1}^{n} \mu_k P[k,i] = \sum_{k=1}^{n} \mu_i Q[i,k] = \mu_i \sum_{k=1}^{n} Q[i,k] = \mu_i.$$

Thus $\underline{\mu}^* = \underline{w}^*$.

Proof

For any i we have

$$\sum_{k=1}^{n} \mu_k P[k,i] = \sum_{k=1}^{n} \mu_i Q[i,k] = \mu_i \sum_{k=1}^{n} Q[i,k] = \mu_i .$$

Thus $\underline{\mu}^* = \underline{w}^*$.

Then by definition $Q = \widehat{P}$.

Kelly's lemma

Corollary

The stationary distribution of the reversed chain $\widehat{P}$ is the same $\underline{w}^*$ as that of the Markov chain P.

Kelly's lemma

Corollary

The stationary distribution of the reversed chain $\widehat{P}$ is the same $\underline{w}^*$ as that of the Markov chain P.

Corollary

Given an irreducible Markov chain P and a probability distribution $\underline{\mu}^*$ such that for all pairs $1 \leq i, k \leq n$ we have $\mu_i P[i, k] = \mu_k P[k, i]$.

Kelly's lemma

Corollary

The stationary distribution of the reversed chain $\widehat{P}$ is the same $\underline{w}^*$ as that of the Markov chain P.

Corollary

Given an irreducible Markov chain P and a probability distribution $\underline{\mu}^*$ such that for all pairs $1 \leq i, k \leq n$ we have $\mu_i P[i,k] = \mu_k P[k,i]$. Then P is reversible, and $\underline{\mu}^* = \underline{w}^*$ is the stationary distribution of P.

Kolmogorov criterion

Kolmogorov criterion

Theorem

An irreducible Markov chain with transition matrix P on n states is reversible iff for all $i_1, \ldots, i_m$ we have

$$P[i_1, i_2]P[i_2 i_3] \cdots P[i_n i_1] = P[i_1, i_n]P[i_n i_{n-1}] \cdots P[i_2 i_1].$$

Kolmogorov criterion

By irreducibility, it clearly follows from the criterion that for all $1 \leq i, k \leq n$ we have $P[i,j] = 0 \Leftrightarrow P[j,i] = 0$.

Kolmogorov criterion

By irreducibility, it clearly follows from the criterion that for all $1 \leq i, k \leq n$ we have $P[i,j] = 0 \Leftrightarrow P[j,i] = 0$.

Hence, it is equivalent to require the criterion to all minimal loops in the digraph corresponding to the Markov chain (whose edges are the pairs with positive transition probability).

Kolmogorov criterion

By irreducibility, it clearly follows from the criterion that for all $1 \leq i, k \leq n$ we have $P[i, j] = 0 \Leftrightarrow P[j, i] = 0$.

Hence, it is equivalent to require the criterion to all minimal loops in the digraph corresponding to the Markov chain (whose edges are the pairs with positive transition probability).

(That weaker requirement also implies that for all $1 \leq i, k \leq n$ we have $P[i, j] = 0 \Leftrightarrow P[j, i] = 0$.)

Proof

Let $\underline{w}^*$ be the stationary distribution of P.

Proof

Let $\underline{w}^*$ be the stationary distribution of P.

NECESSITY OF THE CRITERION

Proof

Let $\underline{w}^*$ be the stationary distribution of P.

NECESSITY OF THE CRITERION

If the chain is reversible, then $w_i P[i,k] = w_k P[k,i]$.

Proof

Let $\underline{w}^*$ be the stationary distribution of P.

NECESSITY OF THE CRITERION

If the chain is reversible, then $w_i P[i,k] = w_k P[k,i]$. Thus

$$\mu_{i_1} \cdots \mu_{i_n} P[i_1, i_2] P[i_2 i_3] \cdots P[i_n i_1] =$$
$$= \mu_{i_2} \cdots \mu_{i_n} \mu_{i_1} P[i_1, i_n] P[i_n i_{n-1}] \cdots P[i_2 i_1].$$

Proof

Let $\underline{w}^*$ be the stationary distribution of P.

NECESSITY OF THE CRITERION

If the chain is reversible, then $w_i P[i,k] = w_k P[k,i]$. Thus

$$\mu_{i_1} \cdots \mu_{i_n} P[i_1, i_2] P[i_2 i_3] \cdots P[i_n i_1] =$$
$$= \mu_{i_2} \cdots \mu_{i_n} \mu_{i_1} P[i_1, i_n] P[i_n i_{n-1}] \cdots P[i_2 i_1].$$

Dividing both sides by $\mu_{i_1} \cdots \mu_{i_n}$ yields the criterion.

Proof

SUFFICIENCY OF THE CRITERION

Proof

Sufficiency of the criterion

Given an $m \in \mathbb{N}$, pair up lists of $m+1$ indices of states starting with i and ending with k, i.

Proof

SUFFICIENCY OF THE CRITERION

Given an $m \in \mathbb{N}$, pair up lists of $m+1$ indices of states starting with i and ending with k, i.

Then the criterion yields $P^m[i,k]P[k,i] = P[i,k]P^m[k,i]$ for all $m \in \mathbb{N}$.

Proof

SUFFICIENCY OF THE CRITERION

Given an $m \in \mathbb{N}$, pair up lists of $m+1$ indices of states starting with i and ending with k, i.

Then the criterion yields $P^m[i,k]P[k,i] = P[i,k]P^m[k,i]$ for all $m \in \mathbb{N}$.

As $m \to \infty$, we obtain $w_k P[k,i] = P[i,k]w_i$,

Proof

SUFFICIENCY OF THE CRITERION

Given an $m \in \mathbb{N}$, pair up lists of $m + 1$ indices of states starting with i and ending with k, i.

Then the criterion yields $P^m[i,k]P[k,i] = P[i,k]P^m[k,i]$ for all $m \in \mathbb{N}$.

As $m \to \infty$, we obtain $w_k P[k,i] = P[i,k]w_i$, which is equivalent to reversibility.

Algorithms, numerical methods

Checking reversibility

Input : the transition matrix P of an irreducible chain.

Checking reversibility

Input : the transition matrix P of an irreducible chain. (If we do not know that it is irreducible, it can be checked in $O(n^2)$ time.)

Checking reversibility

Input : the transition matrix P of an irreducible chain. (If we do not know that it is irreducible, it can be checked in $O(n^2)$ time.)

Question : is the chain reversible ?

Checking reversibility

Input : the transition matrix P of an irreducible chain. (If we do not know that it is irreducible, it can be checked in $O(n^2)$ time.)

Question : is the chain reversible ?

Naive approach :

Checking reversibility

Input : the transition matrix P of an irreducible chain. (If we do not know that it is irreducible, it can be checked in $O(n^2)$ time.)

Question : is the chain reversible ?

Naive approach : compute $\underline{w}^*$, and check the detailed balance equations.

Checking reversibility

Input : the transition matrix P of an irreducible chain. (If we do not know that it is irreducible, it can be checked in $O(n^2)$ time.)

Question : is the chain reversible ?

Naive approach : compute $\underline{w}^*$, and check the detailed balance equations.

This requires the solution of a system of linear equations, hence $\Theta(n^3)$ time with Gaussian elimination, and $\approx \Theta(n^{2.373})$ time with more advanced methods.

Checking reversibility

Input : the transition matrix P of an irreducible chain. (If we do not know that it is irreducible, it can be checked in $O(n^2)$ time.)

Question : is the chain reversible ?

Naive approach : compute $\underline{w}^*$, and check the detailed balance equations.

This requires the solution of a system of linear equations, hence $\Theta(n^3)$ time with Gaussian elimination, and $\approx \Theta(n^{2.373})$ time with more advanced methods. There are also $\Theta(n^2)$ detailed balance equations.

Checking reversibility

Input: the transition matrix P of an irreducible chain. (If we do not know that it is irreducible, it can be checked in $O(n^2)$ time.)

Question: is the chain reversible?

More naive approach:

Input : the transition matrix P of an irreducible chain. (If we do not know that it is irreducible, it can be checked in $O(n^2)$ time.)

Question : is the chain reversible ?

More naive approach : check the Kolmogorov criterion for all minimal loops in the digraph corresponding to the chain.

Checking reversibility

Input: the transition matrix P of an irreducible chain. (If we do not know that it is irreducible, it can be checked in $O(n^2)$ time.)

Question: is the chain reversible?

More naive approach: check the Kolmogorov criterion for all minimal loops in the digraph corresponding to the chain.

That is potentially exponential: there can be an exponential number (in n) of minimal loops; see the exercises.

Better approach :

Checking reversibility

Better approach:

Guess w_1, e.g., $w_1 = 1$, and run a breadth first search to walk through all states.

Better approach:

Guess w_1, e.g., $w_1 = 1$, and run a breadth first search to walk through all states.

Whenever we visit a new state i, we can compute the guesstimate for w_i.

Checking reversibility

Better approach:

Guess w_1, e.g., $w_1 = 1$, and run a breadth first search to walk through all states.

Whenever we visit a new state i, we can compute the guesstimate for w_i.

Normalize the vector obtained to have $\sum w_i = 1$.

Checking reversibility

Better approach:

Guess w_1, e.g., $w_1 = 1$, and run a breadth first search to walk through all states.

Whenever we visit a new state i, we can compute the guesstimate for w_i.

Normalize the vector obtained to have $\sum w_i = 1$.

Check the detailed balance equations with this vector $\underline{w}^*$:

Checking reversibility

Better approach:

Guess w_1, e.g., $w_1 = 1$, and run a breadth first search to walk through all states.

Whenever we visit a new state i, we can compute the guesstimate for w_i.

Normalize the vector obtained to have $\sum w_i = 1$.

Check the detailed balance equations with this vector $\underline{w}^*$: if any of them fails, the chain must be irreversible,

Checking reversibility

Better approach:

Guess w_1, e.g., $w_1 = 1$, and run a breadth first search to walk through all states.

Whenever we visit a new state i, we can compute the guesstimate for w_i.

Normalize the vector obtained to have $\sum w_i = 1$.

Check the detailed balance equations with this vector $\underline{w}^*$: if any of them fails, the chain must be irreversible, if all of them holds, then the chain is reversible with stationary distribution vector $\underline{w}^*$.

Checking reversibility

The runtime is $\Theta(n^2)$.

The runtime is $\Theta(n^2)$.

Recently, Brill, Cheung, Hlynka and Jiang provided a more elegant quadratic algorithm that transforms the transition matrix by operations reminiscent to Gaussian elimination.

Checking reversibility

The runtime is $\Theta(n^2)$.

Recently, Brill, Cheung, Hlynka and Jiang provided a more elegant quadratic algorithm that transforms the transition matrix by operations reminiscent to Gaussian elimination.

The steps preserve reversibility : the end result is symmetric iff the original chain is reversible.

Exercises

Exercises

1. Prove that if P is irreducible, then so is $\widehat{P}$.

Exercises

1. Prove that if P is irreducible, then so is $\widehat{P}$.
2. Show that $\widehat{\widehat{P}} = P$.

Exercises

1. Prove that if P is irreducible, then so is $\widehat{P}$.
2. Show that $\widehat{\widehat{P}} = P$.
3. Prove that if $P = P^*$ then P is reversible. (What is $\underline{w}^*$?)

Exercises

1. Prove that if P is irreducible, then so is $\widehat{P}$.
2. Show that $\widehat{\widehat{P}} = P$.
3. Prove that if $P = P^*$ then P is reversible. (What is $\underline{w}^*$?)
4. Construct a reversible P that is not symmetric with the smallest n possible.

Exercises

1. Prove that if P is irreducible, then so is $\widehat{P}$.
2. Show that $\widehat{\widehat{P}} = P$.
3. Prove that if $P = P^*$ then P is reversible. (What is $\underline{w}^*$?)
4. Construct a reversible P that is not symmetric with the smallest n possible.
5. Let $V = \{x_1, \ldots, x_n, y_1, \ldots, y_n\}$ be the vertices of a graph with edges $x_i x_{i+1}, x_i y_{i+1}, y_i x_{i+1}, y_i y_{i+1}$ (put $x_{n+1} = x_1$ and $y_{n+1} = y_1$). Show that there are at least 1.414^n minimal cycles in this graph.

Exercises

1. Prove that if P is irreducible, then so is $\widehat{P}$.
2. Show that $\widehat{\widehat{P}} = P$.
3. Prove that if $P = P^*$ then P is reversible. (What is $\underline{w}^*$?)
4. Construct a reversible P that is not symmetric with the smallest n possible.
5. Let $V = \{x_1, \dots, x_n, y_1, \dots, y_n\}$ be the vertices of a graph with edges $x_i x_{i+1}, x_i y_{i+1}, y_i x_{i+1}, y_i y_{i+1}$ (put $x_{n+1} = x_1$ and $y_{n+1} = y_1$). Show that there are at least 1.414^n minimal cycles in this graph.
6. Consruct a graph and improve the lower bound to 1.442^n.