Markov chains and applications

Dr. András Pongrácz

Week 7, University of Debrecen

Perron-Frobenius theorem

As we have seen, the matrix $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ shows that the Perron-Frobenius theorem cannot hold for irreducible matrices exactly in its form for primitive matrices.

As we have seen, the matrix $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ shows that the Perron-Frobenius theorem cannot hold for irreducible matrices exactly in its form for primitive matrices.

This has period 3, and indeed, all three thirds roots of unity are eigenvalues (each having modulus 1).

Let $P \in M_n(\mathbb{R})$ be (the transition matrix of) a regular Markov chain, i.e., a primitive stochastic matrix.

Regular case

Let $P \in M_n(\mathbb{R})$ be (the transition matrix of) a regular Markov chain, i.e., a primitive stochastic matrix.

1. The spectral radius $r = \max\{|\lambda| \mid \lambda \in \mathrm{Spec}(P)\} = 1$ is an eigenvalue with multiplicity 1 (Perron root). The corresponding normalized right Perron vector is $\underline{u} = \frac{1}{n}\underline{1}$, and the corresponding normalized left Perron vector $\underline{w}^*$ is also a positive vector, called the stationary distribution.

Regular case

Let $P \in M_n(\mathbb{R})$ be (the transition matrix of) a regular Markov chain, i.e., a primitive stochastic matrix.

1. The spectral radius $r = \max\{|\lambda| \mid \lambda \in \mathrm{Spec}(P)\} = 1$ is an eigenvalue with multiplicity 1 (Perron root). The corresponding normalized right Perron vector is $\underline{u} = \frac{1}{n}\underline{1}$, and the corresponding normalized left Perron vector $\underline{w}^*$ is also a positive vector, called the stationary distribution.

2. There is no other positive (left or right) eigenvector of P than the positive multiples of the Perron vector.

Regular case

Let $P \in M_n(\mathbb{R})$ be (the transition matrix of) a regular Markov chain, i.e., a primitive stochastic matrix.

1. The spectral radius $r = \max\{|\lambda| \mid \lambda \in \mathrm{Spec}(P)\} = 1$ is an eigenvalue with multiplicity 1 (Perron root). The corresponding normalized right Perron vector is $\underline{u} = \frac{1}{n}\underline{1}$, and the corresponding normalized left Perron vector $\underline{w}^*$ is also a positive vector, called the stationary distribution.

2. There is no other positive (left or right) eigenvector of P than the positive multiples of the Perron vector.

3. We have $\lim_{k \to \infty} P^k = \underline{1\,w}^*$.

In particular, starting from any initial probability distribution $\underline{v}$, the probability distribution of the state we are in after k steps $\underline{v}^* P^k$ tends to $\underline{w}^*$.

In particular, starting from any initial probability distribution $\underline{v}$, the probability distribution of the state we are in after k steps $\underline{v}^* P^k$ tends to $\underline{w}^*$. The convergence has exponential speed. (See an earlier exercise.)

In particular, starting from any initial probability distribution $\underline{v}$, the probability distribution of the state we are in after k steps $\underline{v}^* P^k$ tends to $\underline{w}^*$. The convergence has exponential speed. (See an earlier exercise.)

This provides a simple numerical method to approximate the stationary distribution. (Similar idea to Banach's fixed point theorem on the next slide.)

In particular, starting from any initial probability distribution $\underline{v}$, the probability distribution of the state we are in after k steps $\underline{v}^* P^k$ tends to $\underline{w}^*$. The convergence has exponential speed. (See an earlier exercise.)

This provides a simple numerical method to approximate the stationary distribution. (Similar idea to Banach's fixed point theorem on the next slide.)

For an exact approach, we need to solve the system of linear equations $P^* \underline{w} = \underline{w}$.

Banach fixed point theorem

Theorem

Given a complete metric space M and a contraction $P : M \to M$; that is, for some $0 \leq q < 1$, we have $d(Pu, Pv) \leq qd(u, v)$ for ull $u, v \in M$. Then P has a unique fixed point w.

Banach fixed point theorem

Theorem

Given a complete metric space M and a contraction $P : M \to M$; that is, for some $0 \leq q < 1$, we have $d(Pu, Pv) \leq qd(u, v)$ for ull $u, v \in M$. Then P has a unique fixed point w.

Note that $\mathbb{R}^n$ is a complete metric space: it is an elementary result in real analysis that every Cauchy sequence is convergent in $\mathbb{R}^n$ with respect to the standard metric $d(\underline{u}, \underline{v}) = |\underline{u} - \underline{v}| = \sqrt{(\underline{u} - \underline{v})^2}$.

Irreducible case

Irreducible case

Let $P \in M_n(\mathbb{R})$ be (the transition matrix of) an irreducible Markov chain, i.e., an irreducible stochastic matrix.

Irreducible case

Let $P \in M_n(\mathbb{R})$ be (the transition matrix of) an irreducible Markov chain, i.e., an irreducible stochastic matrix. Let h be the period.

Irreducible case

Let $P \in M_n(\mathbb{R})$ be (the transition matrix of) an irreducible Markov chain, i.e., an irreducible stochastic matrix. Let h be the period.

1. The spectral radius $r = \max\{|\lambda| \mid \lambda \in \mathrm{Spec}(P)\} = 1$ is an eigenvalue with multiplicity 1 (Perron root). The corresponding normalized right Perron vector is $\underline{u} = \frac{1}{n}\underline{1}$, and the corresponding normalized left Perron vector $\underline{w}^*$ is also a positive vector, called the stationary distribution.

Irreducible case

Let $P \in M_n(\mathbb{R})$ be (the transition matrix of) an irreducible Markov chain, i.e., an irreducible stochastic matrix. Let h be the period.

1. The spectral radius $r = \max\{|\lambda| \mid \lambda \in \mathrm{Spec}(P)\} = 1$ is an eigenvalue with multiplicity 1 (Perron root). The corresponding normalized right Perron vector is $\underline{u} = \frac{1}{n}\underline{1}$, and the corresponding normalized left Perron vector $\underline{w}^*$ is also a positive vector, called the stationary distribution.
2. There is no other positive (left or right) eigenvector of P than the positive multiples of the Perron vector.

Irreducible case

Let $P \in M_n(\mathbb{R})$ be (the transition matrix of) an irreducible Markov chain, i.e., an irreducible stochastic matrix. Let h be the period.

1. The spectral radius $r = \mathrm{max}\{|\lambda| \mid \lambda \in \mathrm{Spec}(P)\} = 1$ is an eigenvalue with multiplicity 1 (Perron root). The corresponding normalized right Perron vector is $\underline{u} = \frac{1}{n}\underline{1}$, and the corresponding normalized left Perron vector $\underline{w}^*$ is also a positive vector, called the stationary distribution.
2. There is no other positive (left or right) eigenvector of P than the positive multiples of the Perron vector.
3. P has exactly h complex eigenvalues with absolute value 1, namely the h-th roots of unity, and all of them have algebraic multiplicity 1. In fact, the spectrum of P is invariant under multiplication by h-th roots of unity.

Irreducible case

Let $P \in M_n(\mathbb{R})$ be (the transition matrix of) an irreducible Markov chain, i.e., an irreducible stochastic matrix. Let h be the period.

1. The spectral radius $r = \max\{|\lambda| \mid \lambda \in \mathrm{Spec}(P)\} = 1$ is an eigenvalue with multiplicity 1 (Perron root). The corresponding normalized right Perron vector is $\underline{u} = \frac{1}{n}\underline{1}$, and the corresponding normalized left Perron vector $\underline{w}^*$ is also a positive vector, called the stationary distribution.

2. There is no other positive (left or right) eigenvector of P than the positive multiples of the Perron vector.

3. P has exactly h complex eigenvalues with absolute value 1, namely the h-th roots of unity, and all of them have algebraic multiplicity 1. In fact, the spectrum of P is invariant under multiplication by h-th roots of unity.

4. We have $\lim\limits_{k\to\infty} \frac{1}{k} \sum\limits_{i=1}^{k} P^i = \underline{1\,w}^*$.

The case $h = 1$ simplifies to regular matrices.

The case $h = 1$ simplifies to regular matrices. If $h \geq 2$, the transition matrix has a more transparent "canonical" form.

Perron-Frobenius theorem, Irreducible case

The case $h = 1$ simplifies to regular matrices. If $h \geq 2$, the transition matrix has a more transparent "canonical" form. Namely, there is a permutation matrix S such that

$$SAS^{-1} = \begin{pmatrix} 0 & A_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & A_2 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & & \vdots & \vdots \\ \\ 0 & 0 & 0 & 0 & \cdots & 0 & A_{h-1} \\ A_h & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

The zeros along the diagonal are zero square matrices of potentially different sizes, and the matrices $A_1, \ldots, A_h$ are rectangular blocks.

Mean recurrence time

Using the identities in the Perron-Frobenius theorem, it is not hard to show the following basic property of irreducible chains.

Using the identities in the Perron-Frobenius theorem, it is not hard to show the following basic property of irreducible chains.

Theorem

Starting from state s_i in a finite irreducible Markov chain, the mean recurrence time to s_i is $1/w_i$.

Mean recurrence time

Using the identities in the Perron-Frobenius theorem, it is not hard to show the following basic property of irreducible chains.

Theorem

Starting from state s_i in a finite irreducible Markov chain, the mean recurrence time to s_i is $1/w_i$.

Informally, the identity $\lim_{k\to\infty} \frac{1}{k} \sum_{i=1}^{k} P^i = \underline{1w}^*$ shows that we spend w_i proportion of our time in the i-th state in average. The precise proof is an exercise.

Exercises

Exercises

1. Prove the theorem about the main recurrence time.

Exercises

1. Prove the theorem about the main recurrence time.
2. a) Represent coin-toss as an irreducible Markov-chain with two states.
   b) Find the period.
   c) Compute the mean recurrence time and compare it with the general formula.
   d) Compute the second moment of the recurrence time. Make a general conjecture.

Exercises

1. Prove the theorem about the main recurrence time.
2. a) Represent coin-toss as an irreducible Markov-chain with two states.
   b) Find the period.
   c) Compute the mean recurrence time and compare it with the general formula.
   d) Compute the second moment of the recurrence time. Make a general conjecture.
3. Show that if P is irreducible, then $\frac{1}{2}(P + I)$ is regular with the same stationary distribution. (The latter one usually has a better rate of convergence for the iterative method when P has a period greater than 1.)

1. Prove the theorem about the main recurrence time.
2. a) Represent coin-toss as an irreducible Markov-chain with two states.
   b) Find the period.
   c) Compute the mean recurrence time and compare it with the general formula.
   d) Compute the second moment of the recurrence time. Make a general conjecture.
3. Show that if P is irreducible, then $\frac{1}{2}(P + I)$ is regular with the same stationary distribution. (The latter one usually has a better rate of convergence for the iterative method when P has a period greater than 1.)

4. Solve the following exercises to prove the special case of Hoffman's theorem : a graph is bipartite iff its spectrum is symmetric to the origin. (Note that the only if direction is clear.)

Exercises

4. Solve the following exercises to prove the special case of Hoffman's theorem: a graph is bipartite iff its spectrum is symmetric to the origin. (Note that the only if direction is clear.)
   a) Show that an ordered real n-tuple $\lambda_1 \geq \cdots \geq \lambda_n$ is symmetric to the origin iff for all odd $k \in \mathbb{N}$ we have $\sum_{i=1}^{n} \lambda_i^k = 0$.

4. Solve the following exercises to prove the special case of Hoffman's theorem: a graph is bipartite iff its spectrum is symmetric to the origin. (Note that the only if direction is clear.)
   a) Show that an ordered real n-tuple $\lambda_1 \geq \cdots \geq \lambda_n$ is symmetric to the origin iff for all odd $k \in \mathbb{N}$ we have $\sum_{i=1}^{n} \lambda_i^k = 0$.
   b) Observe that if A is the adjacency matrix of a graph G, then $\mathrm{Tr}(A^k)$ is the number of walks of length k with coinciding first and last vertex.

Exercises

4. Solve the following exercises to prove the special case of Hoffman's theorem: a graph is bipartite iff its spectrum is symmetric to the origin. (Note that the only if direction is clear.)
   a) Show that an ordered real n-tuple $\lambda_1 \geq \cdots \geq \lambda_n$ is symmetric to the origin iff for all odd $k \in \mathbb{N}$ we have $\sum_{i=1}^{n} \lambda_i^k = 0$.
   b) Observe that if A is the adjacency matrix of a graph G, then $\mathrm{Tr}(A^k)$ is the number of walks of length k with coinciding first and last vertex.
   c) Prove that $\mathrm{Tr}(A^k) = 0$ iff the main diagonal in A^k is all zero.

4. Solve the following exercises to prove the special case of Hoffman's theorem : a graph is bipartite iff its spectrum is symmetric to the origin. (Note that the only if direction is clear.)
   a) Show that an ordered real n-tuple $\lambda_1 \geq \cdots \geq \lambda_n$ is symmetric to the origin iff for all odd $k \in \mathbb{N}$ we have $\sum_{i=1}^{n} \lambda_i^k = 0$.
   b) Observe that if A is the adjacency matrix of a graph G, then $\mathrm{Tr}(A^k)$ is the number of walks of length k with coinciding first and last vertex.
   c) Prove that $\mathrm{Tr}(A^k) = 0$ iff the main diagonal in A^k is all zero.
   d) Combine these observations to show that the spectrum of A is symmetric to the origin iff there is no walk of odd length in G with coinciding first and last vertex.