Markov Chains and Their Applications

Dr. András Pongrácz

Week 4, University of Debrecen

Perron-Frobenius theorem

Perron-Frobenius theorem

Let $P \in M_n(\mathbb{R})$ be a positive matrix, i.e., a matrix with all entries positive.

Perron-Frobenius theorem

Let $P \in M_n(\mathbb{R})$ be a positive matrix, i.e., a matrix with all entries positive.

1. The spectral radius $r = \max\{|\lambda| \mid \lambda \in \mathrm{Spec}(P)\}$ is an eigenvalue with multiplicity 1 (Perron root), and the corresponding eigenvector $\underline{u}$ (Perron vector) is a positive real vector.

Perron-Frobenius theorem

Let $P \in M_n(\mathbb{R})$ be a positive matrix, i.e., a matrix with all entries positive.

1. The spectral radius $r = \max\{|\lambda| \mid \lambda \in \mathrm{Spec}(P)\}$ is an eigenvalue with multiplicity 1 (Perron root), and the corresponding eigenvector $\underline{u}$ (Perron vector) is a positive real vector.

2. There is no other positive eigenvector of P than the positive multiples of the Perron vector.

Perron-Frobenius theorem

Let $P \in M_n(\mathbb{R})$ be a positive matrix, i.e., a matrix with all entries positive.

1. The spectral radius $r = \max\{|\lambda| \mid \lambda \in \mathrm{Spec}(P)\}$ is an eigenvalue with multiplicity 1 (Perron root), and the corresponding eigenvector $\underline{u}$ (Perron vector) is a positive real vector.
2. There is no other positive eigenvector of P than the positive multiples of the Perron vector.
3. By normalizing $\underline{u}$, that is, picking the constant multiple so that the sum of entries is 1, and similarly picking a normalized Perron vector $\underline{w}$ for P^*, we have $\lim_{k\to\infty} \frac{1}{r^k} P^k = \underline{w}^* \underline{u}$.

Definition

A non-negative matrix $P \in M_n(\mathbb{R})$ is primitive if there is a $k \in \mathbb{N}$ such that P^k is positive.

Generalizations

Definition

A non-negative matrix $P \in M_n(\mathbb{R})$ is primitive if there is a $k \in \mathbb{N}$ such that P^k is positive.

The theorem on the previous slide remains true for primitive matrices without any modification.

Generalizations

Definition

A non-negative matrix $P \in M_n(\mathbb{R})$ is primitive if there is a $k \in \mathbb{N}$ such that P^k is positive.

The theorem on the previous slide remains true for primitive matrices without any modification.

Example: $P = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ is primitive (and not positive).

Definition

A non-negative matrix $P \in M_n(\mathbb{R})$ is primitive if there is a $k \in \mathbb{N}$ such that P^k is positive.

The theorem on the previous slide remains true for primitive matrices without any modification.

Example: $P = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ is primitive (and not positive).

$\chi_P(x) = x^2 - x - 1$ with roots $\frac{1+\sqrt{5}}{2}$, $\frac{1-\sqrt{5}}{2}$, so the Perron root is $r = \frac{1+\sqrt{5}}{2}$.

Generalizations

Definition

A non-negative matrix $P \in M_n(\mathbb{R})$ is primitive if there is a $k \in \mathbb{N}$ such that P^k is positive.

The theorem on the previous slide remains true for primitive matrices without any modification.

Example: $P = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ is primitive (and not positive).

$\chi_P(x) = x^2 - x - 1$ with roots $\frac{1+\sqrt{5}}{2}, \frac{1-\sqrt{5}}{2}$, so the Perron root is $r = \frac{1+\sqrt{5}}{2}$. The corresponding (right) eigenvector is $\underline{u} = (1, \frac{1+\sqrt{5}}{2})^*$.

Irreducible matrices

Definition

Given a non-negative matrix $A \in M_n(\mathbb{R})$, the associated digraph is $G_A = (V, D)$, where V is an n-element set identified by the rows and columns, and $(v_i, v_j) \in D$ iff $A[i, j] > 0$.

Irreducible matrices

Definition

Given a non-negative matrix $A \in M_n(\mathbb{R})$, the associated digraph is $G_A = (V, D)$, where V is an n-element set identified by the rows and columns, and $(v_i, v_j) \in D$ iff $A[i, j] > 0$. Then A is irreducible if the associated matrix is strongly connected.

Irreducible matrices

Definition

Given a non-negative matrix $A \in M_n(\mathbb{R})$, the associated digraph is $G_A = (V, D)$, where V is an n-element set identified by the rows and columns, and $(v_i, v_j) \in D$ iff $A[i, j] > 0$. Then A is irreducible if the associated matrix is strongly connected.

A reducible matrix thus must have a proper subset of vertices $U \subseteq V$ such that there is no edge (u, v) with $u \in U$ and $v \notin U$.

Irreducible matrices

Definition

Given a non-negative matrix $A \in M_n(\mathbb{R})$, the associated digraph is $G_A = (V, D)$, where V is an n-element set identified by the rows and columns, and $(v_i, v_j) \in D$ iff $A[i, j] > 0$. Then A is irreducible if the associated matrix is strongly connected.

A reducible matrix thus must have a proper subset of vertices $U \subseteq V$ such that there is no edge (u, v) with $u \in U$ and $v \notin U$. By rearranging the identification of vertices with rows and columns so that vertices in U come first, we obtain a matrix similar to A (by conjugation with a permutation matrix) of the form $\begin{pmatrix} E & F \\ 0 & G \end{pmatrix}$

Irreducible matrices

Equivalent definitions of irreducible matrices:

Irreducible matrices

Equivalent definitions of irreducible matrices:

1. A non-negative matrix with no proper invariant coordinate subspace, i.e., a subspace spanned by a proper subset of standard basis vectors.

Irreducible matrices

Equivalent definitions of irreducible matrices:

1. A non-negative matrix with no proper invariant coordinate subspace, i.e., a subspace spanned by a proper subset of standard basis vectors.

2. A non-negative matrix A such that there is no permutation matrix P with PAP^{-1} of the form $\begin{pmatrix} E & F \\ 0 & G \end{pmatrix}$.

Irreducible matrices

Equivalent definitions of irreducible matrices:

1. A non-negative matrix with no proper invariant coordinate subspace, i.e., a subspace spanned by a proper subset of standard basis vectors.

2. A non-negative matrix A such that there is no permutation matrix P with PAP^{-1} of the form $\begin{pmatrix} E & F \\ 0 & G \end{pmatrix}$.

Note that in item 1, the condition is about the non-existence of invariant coordinate subspaces, not invariant subspaces in general.

Irreducible matrices

Equivalent definitions of irreducible matrices:

1. A non-negative matrix with no proper invariant coordinate subspace, i.e., a subspace spanned by a proper subset of standard basis vectors.
2. A non-negative matrix A such that there is no permutation matrix P with PAP^{-1} of the form $\begin{pmatrix} E & F \\ 0 & G \end{pmatrix}$.

Note that in item 1, the condition is about the non-existence of invariant coordinate subspaces, not invariant subspaces in general. The latter would be impossible for $n \geq 3$, as every real matrix with $n \geq 3$ has an at most 2-dimensional proper invariant subspace.

Irreducible matrices

Example

$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ is irreducible, but imprimitive.

Irreducible matrices

Example

$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ is irreducible, but imprimitive.

Imprimitivity is witnessed by the fact that powers of A form a periodic sequence with period 3, and all matrices in the period have some zero entries.

Irreducible matrices

Example

$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ is irreducible, but imprimitive.

Imprimitivity is witnessed by the fact that powers of A form a periodic sequence with period 3, and all matrices in the period have some zero entries. And indeed, the Perron-Frobenius theorem no longer holds: the eigenvalues are the third roots of unity, all having the same modulus 1.

Irreducible matrices

Example

$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ is irreducible, but imprimitive.

Imprimitivity is witnessed by the fact that powers of A form a periodic sequence with period 3, and all matrices in the period have some zero entries. And indeed, the Perron-Frobenius theorem no longer holds: the eigenvalues are the third roots of unity, all having the same modulus 1.

This observation leads to the extra property that together with irreducibility guarantees a slightly weaker version of the Perron-Frobenius theorem: see Class 6 for further details.

Gershgorin circles

Gershgorin circles

Definition

Let $A = (a_{i,j}) \in M_n(\mathbb{C})$ be a matrix. Then the n closed discs on the complex plane for all $1 \leq i \leq n$ with center $a_{i,i}$ and radius $\sum_{j \neq i} |a_{i,j}|$ are called Gershgorin discs.

Gershgorin circles

Definition

Let $A = (a_{i,j}) \in M_n(\mathbb{C})$ be a matrix. Then the n closed discs on the complex plane for all $1 \leq i \leq n$ with center $a_{i,i}$ and radius $\sum_{j \neq i} |a_{i,j}|$ are called Gershgorin discs.

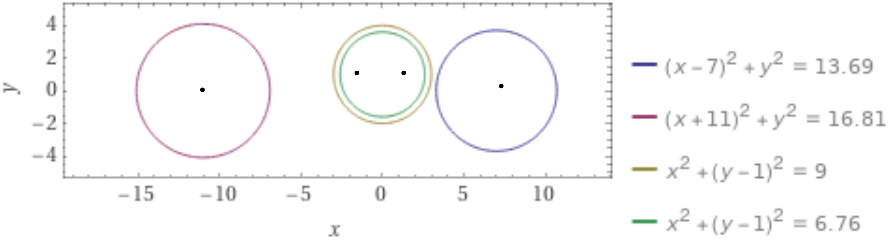
It was observed by Semyon Aronovich Gershgorin that the union of these discs contain the spectrum of the matrix.

Example

$$A = \begin{pmatrix} 7 & 0 & -1.1 & 2.6 \\ 2 & -11 & 0.7 & 1.4 \\ -0.9 & 1.1 & i & 1 \\ 0.7 & -0.3 & 1.6 & i \end{pmatrix}$$

Example

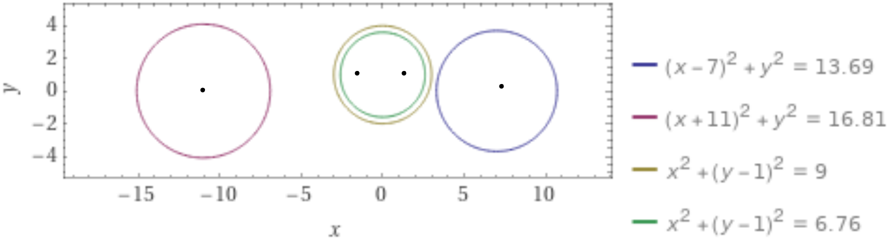
$$A = \begin{pmatrix} 7 & 0 & -1.1 & 2.6 \\ 2 & -11 & 0.7 & 1.4 \\ -0.9 & 1.1 & i & 1 \\ 0.7 & -0.3 & 1.6 & i \end{pmatrix}$$



Spectrum: −11.0402 + 0.0024i, 7.3047+0.0325i, −1.6376 + 0.9662i, 1.3731+0.9988i

Example

$$A = \begin{pmatrix} 7 & 0 & -1.1 & 2.6 \\ 2 & -11 & 0.7 & 1.4 \\ -0.9 & 1.1 & i & 1 \\ 0.7 & -0.3 & 1.6 & i \end{pmatrix}$$



Spectrum: -11.0402 + 0.0024i, 7.3047+0.0325i, -1.6376 + 0.9662i, 1.3731+0.9988i

Note that the union of circles can be divided into three connected components, consisting of (from left to right) one, two and one discs, respectively.

Gershgorin circle theorem

Theorem

Every eigenvalue is contained in at least one Gershgorin disc.

Gershgorin circle theorem

Theorem

Every eigenvalue is contained in at least one Gershgorin disc.

Proof: Let $\underline{u}$ be an eigenvector corresponding to the eigenvalue λ.

Gershgorin circle theorem

Theorem

Every eigenvalue is contained in at least one Gershgorin disc.

Proof: Let $\underline{u}$ be an eigenvector corresponding to the eigenvalue λ. Let $1 \leq i \leq n$ be an index where the modulus of the coordinate $|u[i]|$ is maximal.

Gershgorin circle theorem

Theorem

Every eigenvalue is contained in at least one Gershgorin disc.

Proof: Let $\underline{u}$ be an eigenvector corresponding to the eigenvalue λ. Let $1 \le i \le n$ be an index where the modulus of the coordinate $|u[i]|$ is maximal. As $A\underline{u} = \lambda \underline{u}$, at the i-th coordinate, we obtain $a_{i,i}u[i] + \sum_{j \neq i} a_{i,j}u[j] = \lambda u[i]$, that is, $\sum_{j \neq i} a_{i,j}u[j] = (\lambda - a_{i,i})u[i]$.

Gershgorin circle theorem

Theorem

Every eigenvalue is contained in at least one Gershgorin disc.

Proof: Let $\underline{u}$ be an eigenvector corresponding to the eigenvalue λ. Let $1 \leq i \leq n$ be an index where the modulus of the coordinate $|u[i]|$ is maximal. As $A\underline{u} = \lambda\underline{u}$, at the i-th coordinate, we obtain $a_{i,i}u[i] + \sum_{j \neq i} a_{i,j}u[j] = \lambda u[i]$, that is, $\sum_{j \neq i} a_{i,j}u[j] = (\lambda - a_{i,i})u[i]$.

By taking the modulus of both sides, and using the triangle inequality, we obtain $\sum_{j \neq i} |a_{i,j}||u[j]| \geq |\lambda - a_{i,i}||u[i]|$.

Gershgorin circle theorem

Theorem

Every eigenvalue is contained in at least one Gershgorin disc.

Proof: Let $\underline{u}$ be an eigenvector corresponding to the eigenvalue λ. Let $1 \leq i \leq n$ be an index where the modulus of the coordinate $|u[i]|$ is maximal. As $A\underline{u} = \lambda\underline{u}$, at the i-th coordinate, we obtain $a_{i,i}u[i] + \sum_{j \neq i} a_{i,j}u[j] = \lambda u[i]$, that is, $\sum_{j \neq i} a_{i,j}u[j] = (\lambda - a_{i,i})u[i]$.

By taking the modulus of both sides, and using the triangle inequality, we obtain $\sum_{j \neq i} |a_{i,j}||u[j]| \geq |\lambda - a_{i,i}||u[i]|$. As $|u[j]| \leq |u[i]|$ for all $i \neq j$, we have $\sum_{j \neq i} |a_{i,j}||u[i]| \geq |\lambda - a_{i,i}||u[i]|$,

Gershgorin circle theorem

Theorem

Every eigenvalue is contained in at least one Gershgorin disc.

Proof : Let $\underline{u}$ be an eigenvector corresponding to the eigenvalue λ. Let $1 \leq i \leq n$ be an index where the modulus of the coordinate $|u[i]|$ is maximal. As $A\underline{u} = \lambda\underline{u}$, at the i-th coordinate, we obtain $a_{i,i}u[i] + \sum_{j \neq i} a_{i,j}u[j] = \lambda u[i]$, that is, $\sum_{j \neq i} a_{i,j}u[j] = (\lambda - a_{i,i})u[i]$.

By taking the modulus of both sides, and using the triangle inequality, we obtain $\sum_{j \neq i} |a_{i,j}||u[j]| \geq |\lambda - a_{i,i}||u[i]|$. As $|u[j]| \leq |u[i]|$ for all $i \neq j$, we have $\sum_{j \neq i} |a_{i,j}||u[i]| \geq |\lambda - a_{i,i}||u[i]|$, and then after dividing by $|u[i]|$, the desired inequality $\sum_{j \neq i} |a_{i,j}| \geq |\lambda - a_{i,i}|$ follows.

Eigenvalues on the border

Remark

How can λ be on a circle (the border of a disc) ?

Eigenvalues on the border

Remark

How can λ be on a circle (the border of a disc)? Then in the previous proof, we must have equality in every estimation.

Eigenvalues on the border

Remark

How can λ be on a circle (the border of a disc) ? Then in the previous proof, we must have equality in every estimation. Hence, for all j where $a_{i,j}$ and $u[j]$ are both nonzero, the product $a_{i,j}u[j]$ is $r_j z$ for some complex number z and positive real r_j, and $|u[j]| = |u[i]|$.

Eigenvalues on the border

Remark

How can λ be on a circle (the border of a disc)? Then in the previous proof, we must have equality in every estimation. Hence, for all j where $a_{i,j}$ and $u[j]$ are both nonzero, the product $a_{i,j}u[j]$ is $r_j z$ for some complex number z and positive real r_j, and $|u[j]| = |u[i]|$.

Theorem

If a connected region of the union of discs is the union of k discs, then it contains exactly k eigenvalues with multiplicity (cf. the example).

Proof

Let D be the diagonal matrix obtained by changing all non-diagonal entries of A to 0.

Proof

Let D be the diagonal matrix obtained by changing all non-diagonal entries of A to 0.

Define $M(t) = (1 - t)D + tA$; informally, we are moving on a straight line from D to A with constant speed.

Proof

Let D be the diagonal matrix obtained by changing all non-diagonal entries of A to 0.

Define $M(t) = (1-t)D + tA$; informally, we are moving on a straight line from D to A with constant speed. Indeed, $M(0) = D$ and $M(1) = A$.

Proof

Let D be the diagonal matrix obtained by changing all non-diagonal entries of A to 0.

Define $M(t) = (1 - t)D + tA$; informally, we are moving on a straight line from D to A with constant speed. Indeed, $M(0) = D$ and $M(1) = A$.

Let $(\lambda_1(t), \dots, \lambda_n(t))$ be the labelled spectrum:

Proof

Let D be the diagonal matrix obtained by changing all non-diagonal entries of A to 0.

Define $M(t) = (1 - t)D + tA$; informally, we are moving on a straight line from D to A with constant speed. Indeed, $M(0) = D$ and $M(1) = A$.

Let $(\lambda_1(t), \dots, \lambda_n(t))$ be the labelled spectrum: at time $t = 0$, this is the n-tuple of diagonal entries $(a_{1,1}, \dots, a_{n,n})$ of A (or equivalently, of D).

Proof

Let D be the diagonal matrix obtained by changing all non-diagonal entries of A to 0.

Define $M(t) = (1 - t)D + tA$; informally, we are moving on a straight line from D to A with constant speed. Indeed, $M(0) = D$ and $M(1) = A$.

Let $(\lambda_1(t), \dots, \lambda_n(t))$ be the labelled spectrum: at time $t = 0$, this is the n-tuple of diagonal entries $(a_{1,1}, \dots, a_{n,n})$ of A (or equivalently, of D). Clearly, the labelled spectrum is a continuous function $\mathbb{C} \to \mathbb{C}^n$.

Proof

The Gershgorin discs of $M(t)$ have the same centers for all t, and the radii are linear functions starting from 0 and ending up at the radii of the Gershgorin circles of A.

Proof

The Gershgorin discs of $M(t)$ have the same centers for all t, and the radii are linear functions starting from 0 and ending up at the radii of the Gershgorin circles of A.

Let K be the union of k discs that is disjoint from the union of the other $n - k$ discs K'.

Proof

The Gershgorin discs of $M(t)$ have the same centers for all t, and the radii are linear functions starting from 0 and ending up at the radii of the Gershgorin circles of A.

Let K be the union of k discs that is disjoint from the union of the other $n-k$ discs K'. Then the corresponding k discs for $M(t)$ are contained in K, and the remaining $n-k$ discs are contained in K'.

Proof

The Gershgorin discs of $M(t)$ have the same centers for all t, and the radii are linear functions starting from 0 and ending up at the radii of the Gershgorin circles of A.

Let K be the union of k discs that is disjoint from the union of the other $n-k$ discs K'. Then the corresponding k discs for $M(t)$ are contained in K, and the remaining $n-k$ discs are contained in K'. For $t=0$, these discs contain exactly k eigenvalues.

Proof

The Gershgorin discs of $M(t)$ have the same centers for all t, and the radii are linear functions starting from 0 and ending up at the radii of the Gershgorin circles of A.

Let K be the union of k discs that is disjoint from the union of the other $n - k$ discs K'. Then the corresponding k discs for $M(t)$ are contained in K, and the remaining $n - k$ discs are contained in K'. For $t = 0$, these discs contain exactly k eigenvalues.

By the Gershgorin circle theorem, the spectrum of $M(t)$ is contained in $K \cup K'$ for all t.

Proof

The Gershgorin discs of $M(t)$ have the same centers for all t, and the radii are linear functions starting from 0 and ending up at the radii of the Gershgorin circles of A.

Let K be the union of k discs that is disjoint from the union of the other $n-k$ discs K'. Then the corresponding k discs for $M(t)$ are contained in K, and the remaining $n-k$ discs are contained in K'. For $t=0$, these discs contain exactly k eigenvalues.

By the Gershgorin circle theorem, the spectrum of $M(t)$ is contained in $K \cup K'$ for all t. As K and K' have a positive distance and the labelled spectrum is a continuous function, there are exactly k entries of the labelled spectrum of $M(t)$ in K for all t, and then in particular for $t=1$.

Exercises

Exercises

1. Show that the inverse of a stochastic matrix A is non-negative iff it is stochastic. Prove that if this is the case, then A is a permutation matrix.

Exercises

1. Show that the inverse of a stochastic matrix A is non-negative iff it is stochastic. Prove that if this is the case, then A is a permutation matrix.
2. Generalize the previous exercise to non-negative (invertible) matrices.

Exercises

1. Show that the inverse of a stochastic matrix A is non-negative iff it is stochastic. Prove that if this is the case, then A is a permutation matrix.
2. Generalize the previous exercise to non-negative (invertible) matrices.
3. By using the Jordan normal form of the matrix $P = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$, find the formula for the elements of P^n.

Exercises

1. Show that the inverse of a stochastic matrix A is non-negative iff it is stochastic. Prove that if this is the case, then A is a permutation matrix.
2. Generalize the previous exercise to non-negative (invertible) matrices.
3. By using the Jordan normal form of the matrix $P = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$, find the formula for the elements of P^n.
4. Observe that the series of vectors $\underline{u}_n = (F_n, F_{n+1})^*$, where F_n is the n-th Fibonacci number, satisfies the relations $\underline{u}_0 = (0,1)^*$ and $\underline{u}_n = P\underline{u}_{n-1}$, with P as in the previous exercise. Deduce that $\underline{u}_n = P^n\underline{u}_0$. By the previous exercise, find the exact formula for F_n.

1. Show that the inverse of a stochastic matrix A is non-negative iff it is stochastic. Prove that if this is the case, then A is a permutation matrix.
2. Generalize the previous exercise to non-negative (invertible) matrices.
3. By using the Jordan normal form of the matrix $P = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$, find the formula for the elements of P^n.
4. Observe that the series of vectors $\underline{u}_n = (F_n, F_{n+1})^*$, where F_n is the n-th Fibonacci number, satisfies the relations $\underline{u}_0 = (0,1)^*$ and $\underline{u}_n = P\underline{u}_{n-1}$, with P as in the previous exercise. Deduce that $\underline{u}_n = P^n\underline{u}_0$. By the previous exercise, find the exact formula for F_n.
5. Show that the union of the Gershgorin discs coincide with the spectrum iff the matrix is diagonal.