Stochastic matrices and geometry: a gem from Dmitriev and Dynkin

Ádám BESENYEI

Dept. of Applied Analysis
Eötvös Loránd University, Budapest, Hungary
badam@cs.elte.hu

May 5, 2014

Outline

1. Main problem
   - Setting and some observations
   - History of the problem
2. The paper of Dmitriev and Dynkin from 1945
   - Geometrical reformulation
   - The minimal angle lemma
   - Main results
3. Later results of Dmitriev, Dynkin and Karpelevich
   - Complete solution
   - Generalization
4. Some biographical notes
   - Nikolai Aleksandrovich Dmitriev (1924–2000)
   - Eugene Borisovich Dynkin (1924–)

Main problem

Definition

A matrix $P = (p_{jk})_{j,k=1}^{n}$ is called stochastic if $p_{jk} \geq 0$ for every $j, k = 1, \ldots, n$, i.e., P is nonnegative, and $\sum_{k=1}^{n} p_{jk} = 1$ for every $j = 1, \ldots, n$, i.e, the sum of the entries in each row equals 1.

Main problem

Definition

A matrix $P = (p_{jk})_{j,k=1}^{n}$ is called stochastic if $p_{jk} \geq 0$ for every $j, k = 1, \ldots, n$, i.e., P is nonnegative, and $\sum_{k=1}^{n} p_{jk} = 1$ for every $j = 1, \ldots, n$, i.e, the sum of the entries in each row equals 1.

Question (Kolmogorov, 1938)

Denote

$$M_n := \{\lambda \in \mathbb{C} : \lambda \text{ is an eigenvalue of some } n \times n \text{ stochastic matrix}\}.$$

Determine (or describe) the domain of eigenvalues M_n in the complex plane.

Some observations

Trivial observations

Some observations

Trivial observations

- $1 \in M_n$ for all n since $P\mathbf{1} = \mathbf{1}$ where $\mathbf{1}$ denotes the vector with all entries 1.

Some observations

Trivial observations

- $1 \in M_n$ for all n since $P\mathbf{1} = \mathbf{1}$ where $\mathbf{1}$ denotes the vector with all entries 1.
- M_n is symmetric with respect to the x axis and $M_n \subset \{|z| \leq 1\}$ since

$$\|\lambda v\|_\infty = \|Pv\|_\infty \leq \left(\max_{j=1,\dots,n} \sum_{k=1}^{n} |p_{jk}|\right) \cdot \|v\|_\infty = \|v\|_\infty.$$

  The expression $\|P\|_\infty = \max_{j=1,\dots,n} \sum_{k=1}^{n} |p_{jk}|$ is the so-called row norm of P induced by the vector norm $\|\cdot\|_\infty$.

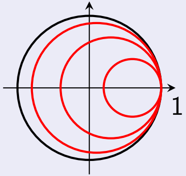
Some observations

Trivial observations

- $1 \in M_n$ for all n since $P\mathbf{1} = \mathbf{1}$ where $\mathbf{1}$ denotes the vector with all entries 1.
- M_n is symmetric with respect to the x axis and $M_n \subset \{|z| \le 1\}$ since

$$\|\lambda v\|_\infty = \|Pv\|_\infty \le \left(\max_{j=1,\dots,n} \sum_{k=1}^{n} |p_{jk}|\right) \cdot \|v\|_\infty = \|v\|_\infty.$$

  The expression $\|P\|_\infty = \max_{j=1,\dots,n} \sum_{k=1}^{n} |p_{jk}|$ is the so-called row norm of P induced by the vector norm $\|\cdot\|_\infty$.
- $M_n \subset \cup_{j=1}^{n} \overline{B(a_{jj}, R_j)}$ where $R_j = \sum_{k \ne j} |a_{jk}| = 1 - a_{jj}$.
  Therefore, $M_n \subset \overline{B(\min a_{jj}, 1 - \min a_{jj})}$.
  These are the so-called Gershgorin discs.
  (Semyon Aronovich Gershgorin (1901–1933))
  (Maurice René Fréchet (1878–1973) ??)



Some observations

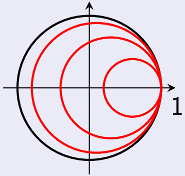
Trivial observations

- $1 \in M_n$ for all n since $P\mathbf{1} = \mathbf{1}$ where $\mathbf{1}$ denotes the vector with all entries 1.
- M_n is symmetric with respect to the x axis and $M_n \subset \{|z| \leq 1\}$ since

$$\|\lambda v\|_\infty = \|Pv\|_\infty \leq \left(\max_{j=1,\dots,n} \sum_{k=1}^{n} |p_{jk}| \right) \cdot \|v\|_\infty = \|v\|_\infty .$$

  The expression $\|P\|_\infty = \max_{j=1,\dots,n} \sum_{k=1}^{n} |p_{jk}|$ is the so-called row norm of P induced by the vector norm $\|\cdot\|_\infty$.

- $M_n \subset \cup_{j=1}^{n} \overline{B(a_{jj}, R_j)}$ where $R_j = \sum_{k \neq j} |a_{jk}| = 1 - a_{jj}$.
  Therefore, $M_n \subset \overline{B(\min a_{jj}, 1 - \min a_{jj})}$.
  These are the so-called Gershgorin discs.
  (Semyon Aronovich Gershgorin (1901–1933))
  (Maurice René Fréchet (1878–1973) ??)

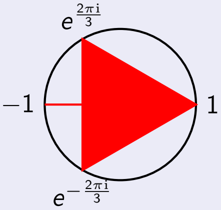


- $M_2 = [-1, 1]$ since the eigenvalues of $P = \begin{bmatrix} p & 1-p \\ q & 1-q \end{bmatrix}$ are 1 and $p - q$.

Some challenging observations

Proposition (try to prove!)

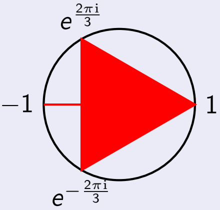
$M_3 = [-1, 1] \cup \Delta$ where Δ is the triangle with vertices $1, \exp(\frac{2\pi i}{3}), \exp(-\frac{2\pi i}{3})$.



Some challenging observations

Proposition (try to prove!)

$M_3 = [-1, 1] \cup \Delta$ where Δ is the triangle with vertices $1, \exp(\frac{2\pi i}{3}), \exp(-\frac{2\pi i}{3})$.



Proposition (try to prove!)

If $|\lambda| = 1$ for some $\lambda \in M_n$, then $\lambda = \exp(2\pi i \frac{p}{q})$ where $0 \leq p \leq q \leq n$ are integers, i.e., λ is a vertex of a regular q-gon inscribed in the unit circle, one of whose vertices is situated at the point 1.

A highly non-trivial observation

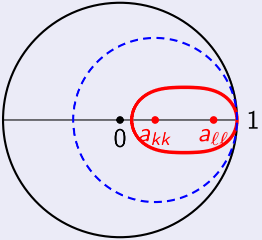
Proposition (try not to prove!)

Let $A = (a_{jk})_{j,k=1}^n$ be a stochastic matrix and a_{kk}, $a_{\ell\ell}$ be the two smallest diagonal entries. Then the eigenvalues of A lie inside or on the boundary of the Cassini oval

$$|z - a_{kk}| \cdot |z - a_{\ell\ell}| \leq (1 - a_{kk}) \cdot (1 - a_{\ell\ell}).$$

If $a_{kk} \neq a_{\ell\ell}$, then this oval lies in the interior of the largest Gershgorin circle, otherwise the two curves coincide.

(This nice result is due to Alfred Theodor Brauer (1894–1985) from 1952.)



History

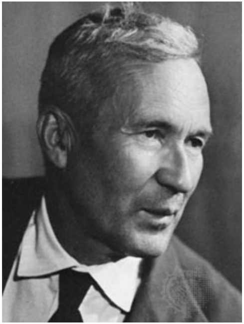
- Vsevolod Ivanovich Romanovsky (1879–1954) a mathematician and statistician, professor at the Central Asia University in Tashkent

History

- Vsevolod Ivanovich Romanovsky (1879–1954) a mathematician and statistician, professor at the Central Asia University in Tashkent
  - 1931 the term stochastic matrix was introduced
  - 1936 survey on stochastic matrices (using results of Oskar Perron (1880–1975) and Ferdinand Georg Frobenius (1849–1917) on nonnegative matrices)

History

- Vsevolod Ivanovich Romanovsky (1879–1954) a mathematician and statistician, professor at the Central Asia University in Tashkent
  - 1931 the term stochastic matrix was introduced
  - 1936 survey on stochastic matrices (using results of Oskar Perron (1880–1975) and Ferdinand Georg Frobenius (1849–1917) on nonnegative matrices)
- Andrey Nikolaevich Kolmogorov (1903–1987)

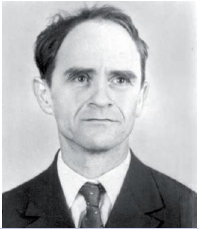


History

- Vsevolod Ivanovich Romanovsky (1879–1954) a mathematician and statistician, professor at the Central Asia University in Tashkent
  - 1931 the term stochastic matrix was introduced
  - 1936 survey on stochastic matrices (using results of Oskar Perron (1880–1975) and Ferdinand Georg Frobenius (1849–1917) on nonnegative matrices)
- Andrey Nikolaevich Kolmogorov (1903–1987)
  - 1938(?) problem of possible location of eigenvalues of stochastic matrices
  - 1944–1945 Moscow State University seminar on Markov chains

History

- Vsevolod Ivanovich Romanovsky (1879–1954) a mathematician and statistician, professor at the Central Asia University in Tashkent
  - 1931 the term stochastic matrix was introduced
  - 1936 survey on stochastic matrices (using results of Oskar Perron (1880–1975) and Ferdinand Georg Frobenius (1849–1917) on nonnegative matrices)
- Andrey Nikolaevich Kolmogorov (1903–1987)
  - 1938(?) problem of possible location of eigenvalues of stochastic matrices
  - 1944–1945 Moscow State University seminar on Markov chains
- Nikolai Aleksandrovich Dmitriev (1924–2000),
  Eugene Borisovich Dynkin (1924–)



History

- Vsevolod Ivanovich Romanovsky (1879–1954) a mathematician and statistician, professor at the Central Asia University in Tashkent
  - 1931 the term stochastic matrix was introduced
  - 1936 survey on stochastic matrices (using results of Oskar Perron (1880–1975) and Ferdinand Georg Frobenius (1849–1917) on nonnegative matrices)
- Andrey Nikolaevich Kolmogorov (1903–1987)
  - 1938(?) problem of possible location of eigenvalues of stochastic matrices
  - 1944–1945 Moscow State University seminar on Markov chains
- Nikolai Aleksandrovich Dmitriev (1924–2000), Eugene Borisovich Dynkin (1924–)
  - 1945 a geometrical reformulation of the problem (both students that time!)
  - 1946 description of M_n for $n \leq 5$

History

- Vsevolod Ivanovich Romanovsky (1879–1954) a mathematician and statistician, professor at the Central Asia University in Tashkent
  - 1931 the term stochastic matrix was introduced
  - 1936 survey on stochastic matrices (using results of Oskar Perron (1880–1975) and Ferdinand Georg Frobenius (1849–1917) on nonnegative matrices)
- Andrey Nikolaevich Kolmogorov (1903–1987)
  - 1938(?) problem of possible location of eigenvalues of stochastic matrices
  - 1944–1945 Moscow State University seminar on Markov chains
- Nikolai Aleksandrovich Dmitriev (1924–2000), Eugene Borisovich Dynkin (1924–)
  - 1945 a geometrical reformulation of the problem (both students that time!)
  - 1946 description of M_n for $n \leq 5$
- Fridrikh Israilevich Karpelevich (1927–2000) (student of Dynkin)

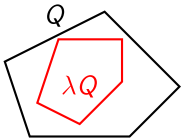
History

- Vsevolod Ivanovich Romanovsky (1879–1954) a mathematician and statistician, professor at the Central Asia University in Tashkent
  - 1931 the term stochastic matrix was introduced
  - 1936 survey on stochastic matrices (using results of Oskar Perron (1880–1975) and Ferdinand Georg Frobenius (1849–1917) on nonnegative matrices)
- Andrey Nikolaevich Kolmogorov (1903–1987)
  - 1938(?) problem of possible location of eigenvalues of stochastic matrices
  - 1944–1945 Moscow State University seminar on Markov chains
- Nikolai Aleksandrovich Dmitriev (1924–2000), Eugene Borisovich Dynkin (1924–)
  - 1945 a geometrical reformulation of the problem (both students that time!)
  - 1946 description of M_n for $n \leq 5$
- Fridrikh Israilevich Karpelevich (1927–2000) (student of Dynkin)
  - 1949, 1951 complete description of the domain M_n

The results of the 1945 paper

Theorem (Geometrical reformulation)

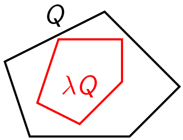
The number $\lambda \in \mathbb{C}$ is an eigenvalue of some $n \times n$ stochastic matrix if and only if there is a convex q-gon Q in the complex plane with number of vertices $q \leq n$ such that $\lambda Q \subset Q$.



The results of the 1945 paper

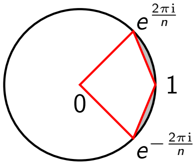
Theorem (Geometrical reformulation)

The number $\lambda \in \mathbb{C}$ is an eigenvalue of some $n \times n$ stochastic matrix if and only if there is a convex q-gon Q in the complex plane with number of vertices $q \leq n$ such that $\lambda Q \subset Q$.



Theorem (Partial description of M_n)

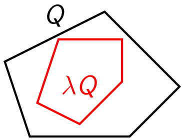
The number $\lambda \in \mathbb{C}$ such that $-\frac{2\pi}{n} \leq \arg\lambda \leq \frac{2\pi}{n}$ is in M_n if and only if it is contained in the quadrilateral whose vertices are the points $0, \exp(-\frac{2\pi i}{n}), 1, \exp(\frac{2\pi i}{n})$.



The results of the 1945 paper

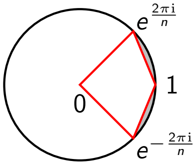
Theorem (Geometrical reformulation)

The number $\lambda \in \mathbb{C}$ is an eigenvalue of some $n \times n$ stochastic matrix if and only if there is a convex q-gon Q in the complex plane with number of vertices $q \le n$ such that $\lambda Q \subset Q$.



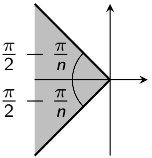
Theorem (Partial description of M_n)

The number $\lambda \in \mathbb{C}$ such that $-\frac{2\pi}{n} \le \arg\lambda \le \frac{2\pi}{n}$ is in M_n if and only if it is contained in the quadrilateral whose vertices are the points $0, \exp(-\frac{2\pi i}{n}), 1, \exp(\frac{2\pi i}{n})$.



Theorem (Eigenvalues of stochastic generators)

The domain of eigenvalues of $n \times n$ matrices $A = (a_{jk})_{j,k=1}^n$ such that $a_{jk} \ge 0$ for $j \ne k$ and $\sum_{k=1}^n a_{jk} = 0$ for all $j = 1, \dots, n$, is the cone $\pi - (\frac{\pi}{2} - \frac{\pi}{n}) \le \arg z \le \pi + (\frac{\pi}{2} - \frac{\pi}{n})$.



The geometrical reformulation

Theorem (Geometrical reformulation)

The number $\lambda \in \mathbb{C}$ is an eigenvalue of some $n \times n$ stochastic matrix if and only if there exists a convex q-gon Q in the complex plane such that $q \leq n$ and $\lambda Q \subset Q$.

The geometrical reformulation

Theorem (Geometrical reformulation)

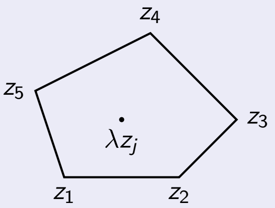
The number $\lambda \in \mathbb{C}$ is an eigenvalue of some $n \times n$ stochastic matrix if and only if there exists a convex q-gon Q in the complex plane such that $q \leq n$ and $\lambda Q \subset Q$.

Proof of the "only if" part.

Let λ be an eigenvalue of $P = (p_{jk})_{j,k=1}^{n}$ and z a corresponding eigenvector. Then

$$\lambda z_j = p_{j1} z_1 + p_{j2} z_2 + \cdots + p_{jn} z_n \quad (j = 1, \ldots, n).$$

Since $p_{jk} \geq 0$ and $p_{j1} + \cdots + p_{jn} = 1$, therefore λz_j lies in the convex hull of the points $z_1, \ldots, z_n$. Let Q be this convex hull, then $\lambda Q \subset Q$.



□

The geometrical reformulation

Theorem (Geometrical reformulation)

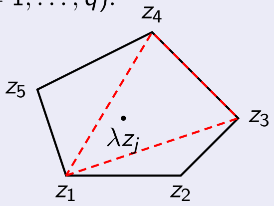
The number $\lambda \in \mathbb{C}$ is an eigenvalue of some $n \times n$ stochastic matrix if and only if there exists a convex q-gon Q in the complex plane such that $q \leq n$ and $\lambda Q \subset Q$.

Proof of the "if" part.

Conversely, if $\lambda Q \subset Q$ for some convex polygon with vertices $z_1, \dots, z_q$, then λz_j lies in Q, thus we may choose $p_{jk} \geq 0$ such that $p_{j1} + \cdots + p_{jn} = 1$ and

$$\lambda z_j = p_{j1} z_1 + p_{j2} z_2 + \cdots + p_{jq} z_q \quad (j = 1, \dots, q).$$

For example, we may choose the barycentric coordinates with respect to some triangle $z_\alpha, z_\beta, z_\gamma$ for $p_{j\alpha}, p_{j\beta}, p_{j\gamma}$, and $p_{jk} = 0$ for other k.



□

Some consequences

Corollary

If $\lambda \in \mathbb{C}$ is an eigenvalue of a stochastic matrix, then $|\lambda| \leq 1$.

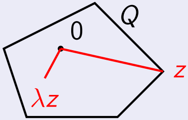
Some consequences

Corollary

If $\lambda \in \mathbb{C}$ is an eigenvalue of a stochastic matrix, then $|\lambda| \leq 1$.

Proof.

Let $z \in Q$ be such that $|z|$ is maximal. Since $\lambda z \in \lambda Q \subset Q$, it follows that $|\lambda z| \leq |z|$, therefore $|\lambda| \leq 1$.



□

Some consequences

Corollary

A number $\lambda \in \mathbb{C}$, $|\lambda| = 1$ is in M_n if and only if $\lambda = \exp(2\pi \mathrm{i} \frac{p}{q})$ where $0 \le p \le q \le n$ are integers.

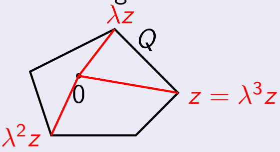
Some consequences

Corollary

A number $\lambda \in \mathbb{C}$, $|\lambda| = 1$ is in M_n if and only if $\lambda = \exp(2\pi i \frac{p}{q})$ where $0 \leq p \leq q \leq n$ are integers.

Proof.

Let again $z \in Q$ be such that $|z|$ is maximal. Then z is some vertex of Q (prove!). Since $\lambda z \in \lambda Q \subset Q$, thus $\lambda z, \lambda^2 z, \ldots$ are also vertices of Q. Consequently, $\lambda^q = 1$ for some $q \leq n$, therefore $\lambda = \exp(2\pi i \frac{p}{q})$, i.e., λ is a vertex of a regular q-gon inscribed in the unit circle one of whose vertices is located at the point 1. But such a regular q-gon multiplied by $\lambda = \exp(2\pi i \frac{p}{q})$ coincides with itself, so $\lambda = \exp(2\pi i \frac{p}{q})$ is indeed an eigenvalue of some stochastic matrix.



□

Some consequences

Corollary

The domain of eigenvalues M_n is star-shaped with respect to the origin.

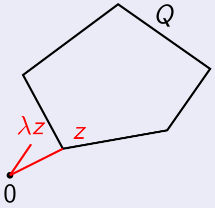
Some consequences

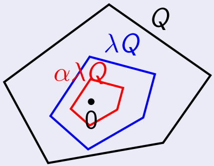
Corollary

The domain of eigenvalues M_n is star-shaped with respect to the origin.

Proof.

If $\lambda \in M_n$, then $\lambda Q \subset Q$ for some Q. We may suppose $0 \in Q$ otherwise there exists a unique $z \in Q$ such that $|z|$ is minimal (prove it!). But then $\lambda z \notin Q$ except the trivial case $\lambda = 1$ when Q can be any polygon. Now since $0 \in Q$, we have for all $0 \leq \alpha \leq 1$ that $(\alpha\lambda)Q \subset \lambda Q \subset Q$, therefore $\alpha\lambda \in M_n$.





□

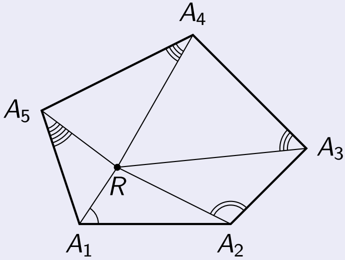
Towards the description of M_n

Lemma (minimal angle)

Let R be an arbitrary point in the interior of a convex n-gon $A_1A_2\dots A_n$ and denote $A_{n+1}=A_1$. Then

$$\min_{k=1,\dots,n} RA_kA_{k+1} \le \frac{\pi}{2}-\frac{\pi}{n}.$$

Equality holds if and only if $A_1A_2\dots A_n$ is a regular n-gon.



Towards the description of M_n

Remark

Special cases of the lemma appeared many times:

- the case $n = 3$ (!) was a problem of the International Mathematical Olympiad in 1991,
- the case $n = 4$ was a problem of a national contest in India in 1991,
- in the years 2000–2001 the Amer. Math. Monthly problems 10824 and 10904 also asked for the cases $n = 3, 4$ and possible generalization, and this latter was marked as a yet unsolved question (but as we see, it was already solved).

Towards the description of M_n

Proof.

We prove by contradiction. Assume that $\sphericalangle RA_kA_{k+1} > \alpha$ for every $k = 1, \ldots, n$ where $\alpha = \pi/2 - \pi/n$ for brevity. Since $\sphericalangle RA_kA_{k+1} > \alpha$, there is a point N_k on the segment RA_{k+1} such that $\sphericalangle RA_kN_k = \alpha$. Denote by ψ_k the angle $\sphericalangle RN_kA_k$. Then the sine theorem in the triangle RA_kN_k implies

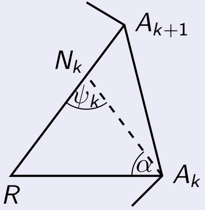
$$\frac{\sin\psi_k}{\sin\alpha} = \frac{RA_k}{RN_k} > \frac{RA_k}{RA_{k+1}}.$$

The product of the above inequality for $k = 1, \ldots, n$ yields

$$\frac{\sin\psi_1}{\sin\alpha} \cdot \frac{\sin\psi_2}{\sin\alpha} \cdot \ldots \cdot \frac{\sin\psi_n}{\sin\alpha} > \frac{RA_1}{RA_2} \cdot \frac{RA_2}{RA_3} \cdot \ldots \cdot \frac{RA_n}{RA_1} = 1,$$

therefore

$$\sin\psi_1 \cdot \ldots \cdot \sin\psi_n > (\sin\alpha)^n.$$



Towards the description of M_n

Proof.

On the other hand

$$\sum_{k=1}^{n} \psi_k = \sum_{k=1}^{n} (\pi - \alpha - \sphericalangle A_k R A_{k+1}) = n(\pi - \alpha) - \sum_{k=1}^{n} \sphericalangle A_k R A_{k+1} = n\alpha.$$

But the product $\sin\psi_1 \cdot \ldots \cdot \sin\psi_n$ in which the angles ψ_k satisfy

$$0 < \psi_1 < \pi, \ldots, 0 < \psi_n < \pi, \quad \sum_{k=1}^{n} \psi_k = n\alpha,$$

attains its maximum for $\psi_1 = \cdots = \psi_n = \alpha$, so

$$\sin\psi_1 \cdot \ldots \cdot \sin\psi_n \le (\sin\alpha)^n.$$

This follows from the inequality of arithmetic and geometric means, combined with Jensen's inequality for the concave sine function on the interval $[0, \pi]$:

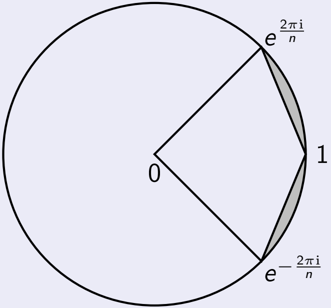
$$\sin\psi_1 \cdot \ldots \cdot \sin\psi_n \le \left(\frac{\sin\psi_1 + \cdots + \sin\psi_n}{n}\right)^n \le \left(\sin\left(\frac{\psi_1 + \cdots + \psi_n}{n}\right)\right)^n = (\sin\alpha)^n.$$

□

Partial description of M_n

Theorem

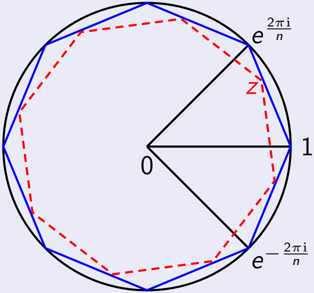
A number $\lambda \in \mathbb{C}$ *such that* $-\frac{2\pi}{n} \leq \arg\lambda \leq \frac{2\pi}{n}$ *is in* M_n *if and only if* λ *is contained in the quadrilateral whose vertices are the points* $0, \exp(-\frac{2\pi i}{n}), 1, \exp(\frac{2\pi i}{n})$.



Partial description of M_n

Proof of the “if” part.

By star-likeness it suffices to show that any point of the segments $[1, e^{\frac{2\pi i}{n}}]$ and $[1, e^{\frac{2\pi i}{n}}]$ may serve as an eigenvalue. Let $z \in [1, e^{\frac{2\pi i}{n}}]$, then the regular n-gon R_n with vertices $1, e^{2\pi i \frac{1}{n}}, \ldots, e^{2\pi i \frac{n-1}{n}}$ multiplied by z is transformed into a polygon all vertices of which are situated on the sides of R_n, therefore $zR_n \subset R_n$.

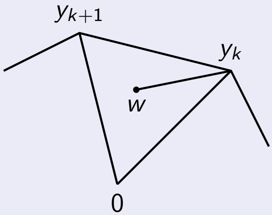


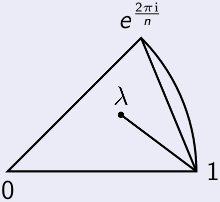
□

Partial description of M_n

Proof of the "only if" part.

Let $\lambda \in M_n$ be such that $0 \le \arg\lambda \le \frac{2\pi}{n}$ and let Y be a convex polygon with vertices $y_1, \dots, y_q$ ($q \le n$) such that $0 \in Y$ and $\lambda Y \subset Y$. By the lemma, there is k such that $\sphericalangle y_{k+1} y_k 0 \le \frac{\pi}{2} - \frac{\pi}{n}$. Since $w := \lambda y_k \in Y$, therefore $\sphericalangle w y_k 0 \le \sphericalangle y_{k+1} y_k 0 \le \frac{\pi}{2} - \frac{\pi}{n}$. But, $\sphericalangle w y_k 0 = \sphericalangle \lambda 1 0$, thus $\sphericalangle \lambda 1 0 \le \frac{\pi}{2} - \frac{\pi}{n}$.





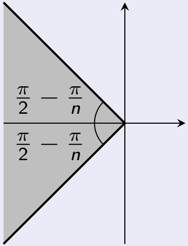
□

Eigenvalues of stochastic generators

Corollary

The domain of eigenvalues of $n \times n$ stochastic generators, i.e., matrices $A = (a_{jk})_{j,k=1}^{n}$ such that $a_{jk} \geq 0$ for $j \neq k$ and $\sum_{k=1}^{n} a_{jk} = 0$ for all $j = 1, \ldots, n$ (in other words, the sum of the entries in each row equals 0), is the cone

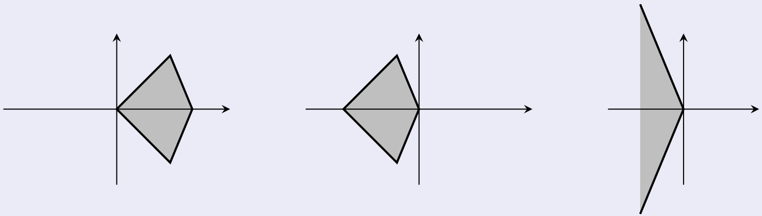
$$\pi - \left(\frac{\pi}{2} - \frac{\pi}{n}\right) \leq \arg z \leq \pi + \left(\frac{\pi}{2} - \frac{\pi}{n}\right).$$



Eigenvalues of stochastic generators

Proof.

If P is a stochastic matrix and $\mu > 0$, then $A = \mu(P - I)$ is a generator, whose eigenvalues are $\mu(\lambda - 1)$. Conversely, if A is a generator, then for $\mu > \max_{j=1,\dots,n} |a_{jj}|$, $P = \frac{1}{\mu}(A + \mu I)$ is a stochastic matrix and $A = \mu(P - I)$.



λ $\lambda - 1$ $\mu(\lambda - 1)$

□

Complete description of M_n for $n \leq 5$

Definition (Dmitriev and Dynkin, 1946)

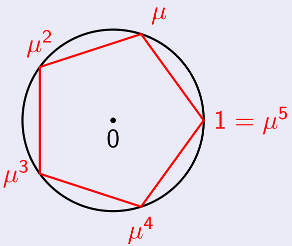
A convex q-gon is called cyclic, generated by μ, if Q is the convex hull of the points $1, \mu, \mu^2, \dots$.

Complete description of M_n for $n \leq 5$

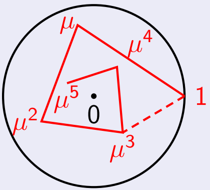
Definition (Dmitriev and Dynkin, 1946)

A convex q-gon is called cyclic, generated by μ, if Q is the convex hull of the points $1, \mu, \mu^2, \ldots$.

Example



$\mu = e^{\frac{2\pi i}{5}}$



$\mu = 0.8e^{\frac{4\pi i}{7}}$

Complete description of M_n for $n \leq 5$

Theorem (Dmitriev and Dynkin, 1946)

For $n \leq 5$, M_n is the union of all cyclic q-gons such that $q \leq n$.

Complete description of M_n for $n \leq 5$

Theorem (Dmitriev and Dynkin, 1946)

For $n \leq 5$, M_n is the union of all cyclic q-gons such that $q \leq n$.

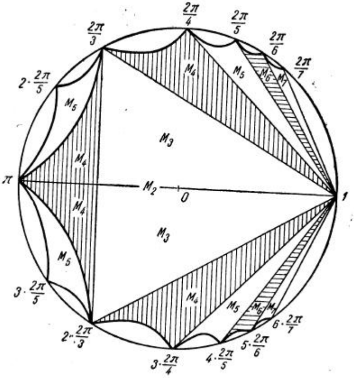
Theorem (Kolmogorov, Dmitriev and Dynkin, 1946)

The eigenvalues of an $n \times n$ nonnegative matrix A lie in the set $\varrho(A) \cdot M_n$ where $\varrho(A) = \max_{j=1,\dots,n} |\lambda_j|$ is the spectral radius of A.

Complete description of M_n for $n \leq 5$

ТЕОРЕМА VII. *Для* $n \leqslant 5$ M_n *будет объединением циклических* k*-угольников* $(k \leqslant n)$.

В заключение приведем чертеж (фиг. 11), схематически изображающий фигуры M_2, M_3, M_4, M_5 и те части фигур M_n для $n > 5$, которые нам известны.



Complete description of M_n

Definition (Karpelevich, 1949, 1951)

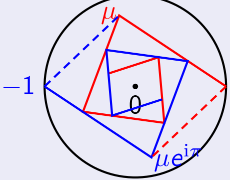
A convex q-gon is called cyclic, if there exist a complex number μ and an integer $p \le q$ such that Q coincides with the convex hull of the system of points $\mu^m e^{\frac{2\pi i r}{p}}$ where $m = 0, 1, \ldots$ and $r = 0, 1, \ldots, p-1$.

Complete description of M_n

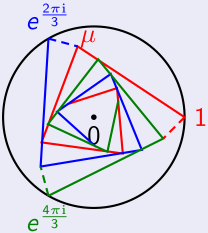
Definition (Karpelevich, 1949, 1951)

A convex q-gon is called cyclic, if there exist a complex number μ and an integer $p \leq q$ such that Q coincides with the convex hull of the system of points $\mu^m e^{\frac{2\pi i r}{p}}$ where $m = 0, 1, \ldots$ and $r = 0, 1, \ldots, p-1$.

Example



$\mu = 0.8e^{\frac{2\pi i}{5}}, p = 2$



$\mu = 0.8e^{\frac{2\pi i}{5}}, p = 3$

Complete description of M_n

Theorem (Karpelevich, 1949, 1951)

The set M_n is the union of all cyclic q-gons such that $q \leq n$.

Complete description of M_n

Theorem (Karpelevich, 1949, 1951)

The set M_n is the union of all cyclic q-gons such that $q \leq n$.

Theorem (Karpelevich, 1951)

The set M_n

- *is symmetric with respect to the origin and contained in the unit circle,*
- *$M_n \cap \{|z| = 1\}$ consists of the points $e^{\frac{2\pi i a}{b}}$ where $0 \leq a < b \leq n$,*
- *∂M_n consists of the previous points and the arcs connecting them in circular order; these arcs can be parametrized by a certain system of equations.*

Extremal eigenvalues

Definition

A number $\lambda \in M_n$ is called an extremal eigenvalue if $\alpha\lambda \notin M_n$ for every $\alpha > 1$.

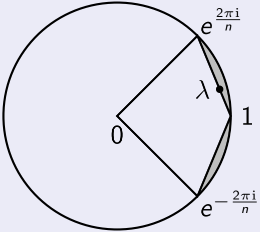
Extremal eigenvalues

Definition

A number $\lambda \in M_n$ is called an extremal eigenvalue if $\alpha\lambda \notin M_n$ for every $\alpha > 1$.

Question

Suppose that a stochastic matrix has an extremal eigenvalue on the segment joining the points 1 and $e^{\frac{2\pi i}{n}}$. What are then the other eigenvalues? How does the matrix look like?



Extremal eigenvalues

Answer

Such a λ can be written in the form

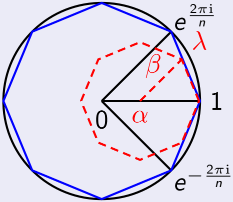
$$\lambda = \alpha + \beta e^{\frac{2\pi i}{n}} \quad \text{where} \quad \alpha, \beta \geq 0, \alpha + \beta = 1.$$

Therefore,

$$(\lambda - \alpha)^n = \beta^n$$

so the characteristic polynomial of the matrix is $(x - \alpha)^n - \beta^n$. Thus, its eigenvalues are $\lambda_j = \alpha + \beta\varepsilon_j$ where the ε_j are the nth roots of unity (i.e., the eigenvalues are vertices of a regular n-gon). A stochastic matrix with such eigenvalues is

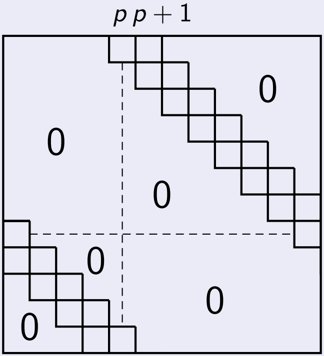
$$\begin{bmatrix} \alpha & \beta & 0 & 0 & 0 \\ 0 & \alpha & \beta & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \alpha & \beta \\ \beta & 0 & 0 & 0 & \alpha \end{bmatrix}.$$



Extremal eigenvalues

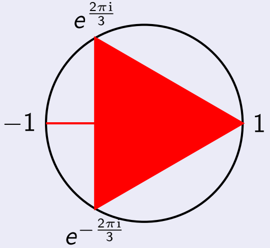
Theorem (Dmitriev and Dynkin, 1946)

If an $n \times n$ stochastic matrix has an extremal eigenvalue λ such that $\lambda \in M_n \setminus M_{n-1}$ and $\frac{2\pi p}{n} \leq \arg \lambda \leq \frac{2\pi(p+1)}{n}$ where $0 \leq p \leq n-1$, then the matrix can be transformed to the following schematic form



Extremal eigenvalues

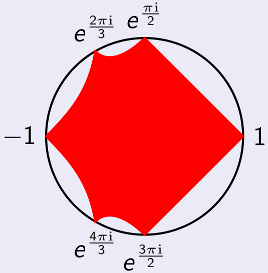
Case of M_3



Angle	Equation of extremal roots (where $\alpha, \beta \geq 0, \alpha + \beta = 1$)
$(-\frac{2\pi}{3}, \frac{2\pi}{3})$	$(\lambda - \alpha)^3 = \beta^3$
$(\frac{2\pi}{3}, \frac{4\pi}{3})$	$\lambda^3 = \alpha + \beta\lambda$ (prove!)

Extremal eigenvalues

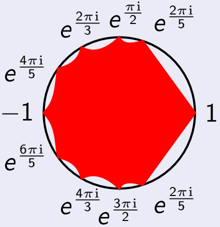
Case of M_4



Angle	Equation of extremal roots (where $\alpha, \beta \geq 0, \alpha + \beta = 1$)
$(-\frac{2\pi}{4}, \frac{2\pi}{4})$	$(\lambda - \alpha)^4 = \beta^4$
$(\frac{\pi}{2}, \frac{2\pi}{3}), (\frac{4\pi}{3}, \frac{3\pi}{2})$	$\lambda^4 = \alpha + \beta\lambda$
$(\frac{2\pi}{3}, \frac{4\pi}{3})$	$(\lambda^2 - \alpha)^2 = \beta^2\lambda$

Extremal eigenvalues

Case of M_5



Angle	Equation of extremal roots (where $\alpha, \beta \geq 0, \alpha + \beta = 1$)
$(-\frac{2\pi}{5}, \frac{2\pi}{5})$	$(\lambda - \alpha)^5 = \beta^5$
$(\frac{2\pi}{5}, \frac{\pi}{2}), (\frac{3\pi}{2}, \frac{8\pi}{5})$	$\lambda^5 = \alpha + \beta\lambda$
$(\frac{\pi}{2}, \frac{2\pi}{3}), (\frac{4\pi}{3}, \frac{3\pi}{2})$	$\lambda^4 = \alpha + \beta\lambda$
$(\frac{2\pi}{3}, \frac{4\pi}{5}), (\frac{6\pi}{5}, \frac{4\pi}{3})$	$\lambda^5 = \alpha + \beta\lambda^2$
$(\frac{4\pi}{5}, \frac{6\pi}{5})$	$\lambda(\lambda^2 - \alpha)^2 = \beta^2$

Nonnegative matrices

Theorem (Kolmogorov, Dmitriev and Dynkin, 1946)

The eigenvalues of an $n \times n$ nonnegative matrix A lie in the set $\varrho(A) \cdot M_n$ where $\varrho(A) = \max_{j=1,\dots,n} |\lambda_j|$ is the spectral radius of A.

Nonnegative matrices

Theorem (Kolmogorov, Dmitriev and Dynkin, 1946)

The eigenvalues of an $n \times n$ nonnegative matrix A lie in the set $\varrho(A) \cdot M_n$ where $\varrho(A) = \max_{j=1,\dots,n} |\lambda_j|$ is the spectral radius of A.

Proof.

If A is irreducible, i.e., it is not similar to a block triangular matrix, then by Perron's results there is a positive eigenvector $x \in \mathbb{R}^n$ such that $Ax = \varrho x$. Let X be the diagonal matrix with diagonal entries $X_{jj} = x_j$, then $B := \frac{1}{\varrho} X^{-1} A X$ is a stochastic matrix, since

$$X^{-1} A X \mathbf{1} = X^{-1} A x = X^{-1} \varrho x = \varrho \mathbf{1}.$$

The eigenvalues of B are λ_j / ϱ which lie in M_n, therefore $\lambda_j \in \varrho(A) \cdot M_n$.

Nonnegative matrices

Theorem (Kolmogorov, Dmitriev and Dynkin, 1946)

The eigenvalues of an $n \times n$ nonnegative matrix A lie in the set $\varrho(A) \cdot M_n$ where $\varrho(A) = \max_{j=1,\dots,n} |\lambda_j|$ is the spectral radius of A.

Proof.

If A is reducible, then A is similar to a block triangular matrix

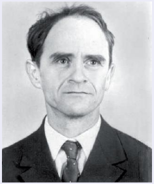
$$\begin{bmatrix} A_{11} & 0 & 0 & \dots & 0 \\ A_{21} & A_{22} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ A_{k1} & A_{k2} & A_{k3} & \dots & A_{kk} \end{bmatrix}$$

where $A_{11}, \dots, A_{kk}$ are irreducible. Each eigenvalue of A is an eigenvalue of some A_{jj}, therefore they lie in the set $\varrho(A_{jj}) \cdot M_n \subset \varrho(A) \cdot M_n$. □

Some biographical notes

Nikolai Aleksandrovich Dmitriev (1924–2000)

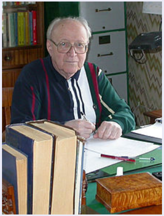
- newspapers: young Kolya is a *"phenomenon that appears once in a century"*,
- enrolled in Moscow State University at the age of 14,
- later he was working at the Russian Federal Nuclear Center,
- Kolmogorov: *"Why do you need those computers? You have Kolya Dmitriev, don't you?"*,
- Andrei Sakharov (1921–1989, Nobel Peace Prize Laureate): *"...perhaps the only one among us with the sparks of God. You could think that Kolya is this quiet, modest boy. But we all tremble before him, as if he were the highest judge."*



Some biographical notes

Eugene Borisovich Dynkin (1924–)

- enrolled in Moscow State University at the age of 16,
- 1977–2012 professor at Cornell University,
- fundamental contributions to algebra and probability theory,
- The Eugene B. Dynkin Collection of Mathematics Interviews: `http://dynkincollection.library.cornell.edu/`
- Seventy years in mathematics: from 28:50 Dynkin recounts the problem `http://www.youtube.com/watch?v=IW3QRmSviMI`



References

N. Dmitriev, E. Dynkin, On the characteristic roots of stochastic matrices, *C. R. (Doklady) Acad. Sci. URSS*, **49** (1945), 159–162.

N. Dmitriev, E. Dynkin, On characteristic roots of stochastic matrices *Izv. Akad. Nauk SSSR Ser. Mat.*, **10** (1946), 167-184.

F. I. Karpelevich, On the characteristic roots of matrices with nonnegative elements, *Izv. Akad. Nauk SSSR Ser. Mat.*, **15** (1951), 361-383.

V. Romanovsky, Sur les zéros de matrices stochastiques, *C. R. Acad. Sc., Paris*, **192** 1931, 267–269.

V. Romanovsky, Recherches sur les chaînes de Markoff, *Acta Math.*, **66** 1936, 147–251.

Joanne Swift, *The location of characteristic roots of stochastic matrices*, MSc Thesis, McGill University, 1972.

V. S. Vladimirov et al, Nikolai Aleksandrovich Dmitriev (obituary), *Russ. Math. Surv.*, **56** (2001), 403–406.

The End

Thank you for your attention!