Joint spectral radius

Constrained matrix products☆

Victor Kozyakin
kozyakin@iitp.ru

Institute for Information Transmission Problems
(Kharkevich Institute)
Russian Academy of Sciences

EQINOCS (entropy and information in computational systems) workshop
Paris Diderot University

May 9–11, 2016

☆ Supported by the Russian Science Foundation, Project No. 14–50–00150.

Outline

Joint Spectral Radius

Markovian Matrix Products

Frequency Constrained Matrix Products

Joint Spectral Radius

Let $\mathcal{A} = \{A_1, A_2, \ldots, A_r\}$ be a set of $(d \times d)$-matrices.

When the matrix products $A_{i_n} \cdots A_{i_2} A_{i_1}$ converge/diverge ?

- “parallel” vs “sequential” computations (e.g., Gauss-Seidel vs Jacobi method, distributed computations);
- “asynchronous” vs “synchronous” data exchange (control theory, large-scale networks);
- smoothness of Daubechies wavelets (computational mathematics);
- one-dimensional discrete Schrödinger equations with quasiperiodic potentials (theory of quasicrystals);
- affine iterated function systems (theory of fractals);
- Hopfield-Tank neural networks (biology, mathematics);
- “triangular arbitrage” (market economics);
- etc.

Rota–Strang Formula

Let $\|\cdot\|$ be a sub-multilicative matrix norm, i.e. $\|AB\| \le \|A\|\cdot\|B\|$ for any matrices A, B. Define a generalization of the quantity $\|A^n\|$ to the case of several matrices:

$$\rho_n(\mathcal{A}) = \max_{A_{i_j} \in \mathcal{A}} \|A_{i_n} \cdots A_{i_1}\|, \qquad n \ge 1.$$

Definition (Rota & Strang, 1960)

$$\rho(\mathcal{A}) := \limsup_{n\to\infty} \rho_n(\mathcal{A})^{1/n} \quad \left(= \inf_{n\ge1} \rho_n(\mathcal{A})^{1/n}\right),$$

is called the ***joint spectral radius (JSR)*** of $\mathcal{A}$.

Remark

$\rho(\mathcal{A})$ does not depend on the norm $\|\cdot\|$.

Daubechies–Lagarias Formula

Similarly define a generalization of the quantity $\rho(A^n) = \rho(A)^n$ to the case of several matrices:

$$\bar{\rho}_n(\mathcal{A}) = \max_{A_{i_j} \in \mathcal{A}} \rho(A_{i_n} \cdots A_{i_1}), \qquad n \geq 1.$$

DEFINITION (DAUBECHIES & LAGARIAS, 1992)

$$\bar{\rho}(\mathcal{A}) := \limsup_{n\to\infty} \bar{\rho}_n(\mathcal{A})^{1/n} \quad \left(= \sup_{n\geq 1} \bar{\rho}_n(\mathcal{A})^{1/n} \right),$$

is called the ***generalized spectral radius (GSR)*** of $\mathcal{A}$.

Berger–Wang Formula

Theorem (Berger & Wang, 1992)

If the set $\mathcal{A}$ is bounded then GSR=JSR:

$$\bar{\rho}(\mathcal{A}) = \rho(\mathcal{A}).$$

This theorem is of crucial importance in numerous constructions of the theory of joint/generalized spectral radius.

Most computational methods of evaluating JSR/GSR are based on the following

Corollary

$$\bar{\rho}_n(\mathcal{A})^{1/n} \le \bar{\rho}(\mathcal{A}) = \rho(\mathcal{A}) \le \rho_n(\mathcal{A})^{1/n}, \qquad \forall n.$$

Alternative Formulae for JSR/GSR

- *Elsner, 1995; Shih, 1999* — via infimum of norms;
- *Chen & Zhou, 2000* — via trace of matrix products;
- *Parrilo & Jadbabaie, 2008* — via homogeneous polynomials instead of norms;
- *Blondel & Nesterov, 2005* — via Kronecker (tensor) products of matrices;
- *Barabanov, 1988; Protasov, 1996* — via special kind of norms with additional properties;
- etc.

Lower Spectral Radius

Let again $\|\cdot\|$ be a sub-multilicative matrix norm. Define

$$\check{\rho}_n(\mathcal{A}) = \min_{A_{i_j}\in\mathcal{A}} \|A_{i_n}\cdots A_{i_1}\|, \qquad n \geq 1.$$

Definition (Gurvits, 1995)

$$\check{\rho}(\mathcal{A}) := \lim_{n\to\infty} \check{\rho}_n(\mathcal{A})^{1/n} \quad \left(= \inf_{n\geq 1} \check{\rho}_n(\mathcal{A})^{1/n}\right),$$

is the ***lower spectral radius (LSR)*** of $\mathcal{A}$.

Difference between LSR and JSR:

- $\rho(\mathcal{A}) < 1 \quad \Longrightarrow \quad$ ***stability*** of $\mathcal{A}$;
- $\check{\rho}(\mathcal{A}) < 1 \quad \Longrightarrow \quad$ ***stabilizability*** of $\mathcal{A}$.

Lower Spectral Radius (cont.)

- LSR possesses “less stable” continuity properties than JSR, see Bousch & Mairesse, 2002;
- Until recently, “good” properties of the LSR, including numerical algorithms of computation, were obtained **only** for matrix sets $\mathscr{A}$ having an invariant cone, see Protasov, Jungers & Blondel, 2009/10; Jungers, 2012; Guglielmi & Protasov, 2013;
- Bochi & Morris, 2015, started a systematic investigation of the continuity properties of the LSR, giving in particular a ***sufficient condition for Lipschitz continuity of the LSR***.

  Their investigation is based on the concepts of *dominated splitting* and *k-multicones* from the theory of hyperbolic linear cocycles.

Recent Trends

Number of publications since 1960 so far, directly related to the JSR/GSR theory, totals about 360, see, e.g. Kozyakin, 2013.

More than 100 publications in the last five years

Most important (*to my mind* !) directions:

- Numerical algorithms for computation of the JSR;
- Investigation of the LSR;
- Measure theoretic and ergodic methods.

Numerical Algorithms

- *Maesumi, 1996; Gripenberg, 1996*: branch-and-bound methods based on the formula
$$\bar{\rho}_n(\mathcal{A})^{1/n} \le \bar{\rho}(\mathcal{A}) = \rho(\mathcal{A}) \le \rho_n(\mathcal{A})^{1/n};$$
- *Blondel & Nesterov, 2005*: algorithms based on the formula
$$\rho(\mathcal{A}) = \lim_{k\to\infty} \rho^{1/k}(A_1^{\otimes k} + \cdots + A_m^{\otimes k})$$
expressing the JSR of matrices with non-negative entries via Kronecker powers of the matrices $A_i \in \mathcal{A}$;
- *Nesterov, 2000; Parrilo, 2000; Parrilo & Jadbabaie, 2007; Legat, Jungers & Parrilo, 2016; etc.*: approximation of the JSR using the sum of squares (SoS) techniques;
- *Guglielmi & Zennaro, 2005; Guglielmi & Protasov, 2013; Protasov, 2016*: approximation of the JSR by constructing polygon approximation of extremal norms;
- *Kozyakin, 2010; Kozyakin, 2011*: relaxation algorithms for iterative building of Barabanov norms and computation of the JSR.

MATLAB® toolboxes

JSR toolbox (combines 7 different algorithms):

Vankeerberghen, Hendrickx, Jungers, Chang & Blondel, 2011;

Chang & Blondel, 2013;

Vankeerberghen, Hendrickx & Jungers, 2014

Joint spectral radius computation toolbox:

Protasov & Jungers, 2012;

Cicone & Protasov, 2012;

Guglielmi & Protasov, 2013

Measure Theoretic and Ergodic Methods

Ideas of the measure and ergodic theory underlie various facts of the theory of JSR/GSR, see

Neumann & Schneider, 1999;

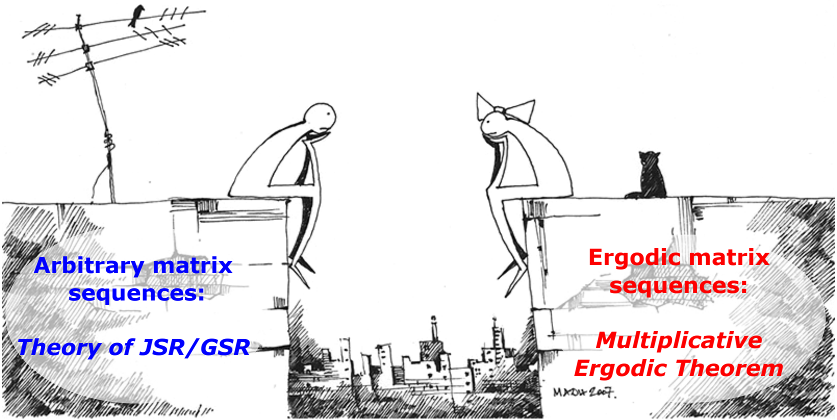
Bousch & Mairesse, 2002;

Morris, 2010; Morris, 2012; Morris, 2013;

Dai, Huang & Xiao, 2008; Dai, Huang & Xiao, 2011a;
Dai, Huang & Xiao, 2011b; Dai, Huang & Xiao, 2013;

Dai, 2011; Dai, 2012; Dai, 2014;

etc.



What is in between?

In the evening of the day. Sergei Ivanov, 2016

An Illusive Bridge

Let Σ_K^+ be the space of infinite sequences $\sigma : \mathbb{N} \to \{1, 2, \ldots, K\}$ endowed with the product topology, and let θ be the Markov (or Bernoulli) shift on Σ_K^+:

$$\theta : \{i_1, i_2, i_3, \ldots\} \mapsto \{i_2, i_3, i_4, \ldots\}.$$

A Borel measure μ on Σ_K^+ is called ***ergodic*** if it is θ-invariant and $\mu\left(S \triangle \theta^{-1}(S)\right) = 0$ implies $\mu(S) = 0$ or $\mu(S) = 1$.

Theorem (Dai, Huang & Xiao, 2011b)

Given a finite set of matrices $\mathcal{A} \subset \mathbb{C}^{d\times d}$, there exists an ergodic Borel probability measure μ_* on Σ_K^+ such that

$$\rho(\mathcal{A}) = \lim_{n\to\infty} \|A_{i_1} A_{i_2} \cdots A_{i_n}\|^{1/n}, \quad A_{i_j} \in \mathcal{A},$$

for μ_*-a.e. sequences $\{i_1, i_2, \ldots, i_n, \ldots\}$.

An Illusive Bridge (cont.)

Remark

One should keep in mind that the sequences which are realized ***almost everywhere*** in some shift-invariant Borel measure may be rather “lean” from the “common point of view”.

Markovian Matrix Products

Markovian Matrix Products

Given: a set of matrices $\mathcal{A} = \{A_1, A_2, \ldots, A_r\}$ and
an $(r \times r)$-matrix $\Omega = (\omega_{ij}):\quad \omega_{ij} \in \{0, 1\}$.

Definition

The matrix product $A_{i_n} \cdots A_{i_1}$ is called ***Markovian*** if each pair of indices $\{i_k, i_{k+1}\}$ is ***Ω-admissible***, i.e.

$$\omega_{i_{k+1} i_k} \equiv 1, \qquad k = 1, 2, \ldots, n-1.$$

Remark

The question on existence of infinite Ω-admissible sequences $\{i_k\}$ is decidable algorithmically in a finite number of steps.

Markovian Joint/Generalized Spectral Radii

In particular, if each column of the transition matrix Ω is non-zero then the following quantities are defined for any n:

$$\rho_n(\mathcal{A},\Omega) := \max\left\{\|A_{i_n}\cdots A_{i_1}\| \;:\; \omega_{i_{k+1}i_k} = 1 \text{ for all } k\right\},$$
$$\bar{\rho}_n(\mathcal{A},\Omega) := \max\left\{\rho(A_{i_n}\cdots A_{i_1}) \;:\; \omega_{i_{k+1}i_k} = 1 \text{ for all } k\right\}.$$

Definition

$$\rho(\mathcal{A},\Omega) := \limsup_{n\to\infty} \rho_n(\mathcal{A},\Omega)^{1/n},$$
$$\bar{\rho}(\mathcal{A},\Omega) := \limsup_{n\to\infty} \bar{\rho}_n(\mathcal{A},\Omega)^{1/n}$$

are called the ***Markovian joint/generalized spectral radii*** of $\mathcal{A}$.

Dai Theorem

The Markovian spectral radius was first (?) introduced by *Dai, 2012* under the name ***spectral radius with constraints***.

Nowadays, in the case when the matrix sequences are generated by finite automata, the term ***constrained spectral radius*** is sometimes used, see Philippe, Essick, Dullerud & Jungers, 2015; Philippe & Jungers, 2015; Legat, Jungers & Parrilo, 2016.

Theorem (Dai, 2012; Dai, 2014)

$$\bar{\rho}^{(per)}(\mathcal{A},\Omega)=\bar{\rho}(\mathcal{A},\Omega)=\rho(\mathcal{A},\Omega),$$

where $\bar{\rho}^{(per)}(\mathcal{A},\Omega)$ is obtained by restricting of $\bar{\rho}(\mathcal{A},\Omega)$ to the ***periodic*** *Markovian products of matrices.*

Dai Theorem (cont.)

REMARK

So far all known proofs of the Berger-Wang formula relied on the arbitrariness of matrix products involved

$$\Downarrow$$

Dai's generalization of the Berger-Wang formula is nontrivial and difficult.

The original proof of Dai's theorem was based on a ponderous machinery of ergodic theory.

Below, we describe a much simpler approach suggested in Kozyakin, 2014a.

Ω-lifting Techniques

Given a set of $(d \times d)$-matrices $\mathcal{A} = \{A_1, A_2, \ldots, A_r\}$ and a *transition matrix* $\boldsymbol{\Omega} = (\omega_{ij})_{i,j=1}^{r}$, define matrices

$$\Omega_i = \boldsymbol{\omega}_i^T \boldsymbol{\delta}_i, \quad A^{(i)} := \Omega_i \otimes A_i, \qquad i = 1, 2, \ldots, r,$$

where $\boldsymbol{\omega}_i = \{\omega_{1i}, \ldots, \omega_{ri}\}$, $\boldsymbol{\delta}_i = \{\delta_{1i}, \ldots, \delta_{ri}\}$, δ_{ij} is the Kronecker symbol, and $\otimes$ is the Kronecker product of matrices.

Definition

The set of matrices $\mathcal{A}_\Omega := \{A^{(i)}\}$ is called the ***$\boldsymbol{\Omega}$-lift*** of the set of matrices $\mathcal{A}$.

Example

Let

$$\Omega = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Then

$$\Omega_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \Omega_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \Omega_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

$$A^{(1)} = \begin{pmatrix} A_1 & 0 & 0 \\ 0 & 0 & 0 \\ A_1 & 0 & 0 \end{pmatrix}, \quad A^{(2)} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & A_2 & 0 \end{pmatrix}, \quad A^{(3)} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & A_3 \\ 0 & 0 & 0 \end{pmatrix}.$$

Crucial Observation

Let $\|\cdot\|$ be a sub-multiplicative norm on the space of $(d \times d)$-matrices.

Then, for a block $(r \times r)$-matrix $M = (m_{ij})$ with the $(d \times d)$-matrix elements m_{ij}, define the norm

$$|\!|\!|M|\!|\!| := \max_{1 \le i \le r} \sum_{j=1}^{r} \|m_{ij}\|.$$

The norm $|\!|\!|\cdot|\!|\!|$ is also sub-multiplicative.

Then

$$\|A^{(i_n)}\cdots A^{(i_1)}\| = \begin{cases} \|A_{i_n}\cdots A_{i_1}\| & \text{if } \omega_{i_{k+1}i_k} \equiv 1; \\ 0 & \text{in the opposite case.} \end{cases}$$

$$\Downarrow$$

$$\rho_n(\mathcal{A}_\Omega) = \rho_n(\mathcal{A}, \Omega) \qquad \forall\, n.$$

$$\Downarrow$$

$$\rho(\mathcal{A}_\Omega) = \rho(\mathcal{A}, \Omega).$$

Crucial Observation (cont.)

Similarly

$$\rho(A^{(i_n)}\cdots A^{(i_1)}) = \begin{cases} \rho(A_{i_n}\cdots A_{i_1}) & \text{if } \omega_{i_{k+1}i_k}\equiv 1 \text{ and } \omega_{i_1 i_n}=1; \\ 0 & \text{in the opposite case.} \end{cases}$$

$$\Downarrow$$

$$\bar{\rho}_n(\mathcal{A}_\Omega) = \bar{\rho}_n^{(\mathrm{per})}(\mathcal{A},\Omega) \qquad \forall\, n.$$

$$\Downarrow$$

$$\bar{\rho}(\mathcal{A}_\Omega) = \bar{\rho}^{(\mathrm{per})}(\mathcal{A},\Omega).$$

Proof of Dai's Theorem

By the Berger–Wang theorem

$$\bar{\rho}(\mathcal{A}_{\Omega}) = \rho(\mathcal{A}_{\Omega}).$$

Then by the earlier made observations

$$\bar{\rho}^{(\mathrm{per})}(\mathcal{A}, \Omega) = \bar{\rho}(\mathcal{A}_{\Omega}) = \rho(\mathcal{A}_{\Omega}) = \rho(\mathcal{A}, \Omega),$$

from which Dai's theorem immediately follows:

$$\bar{\rho}^{(\mathrm{per})}(\mathcal{A}, \Omega) = \bar{\rho}(\mathcal{A}, \Omega) = \rho(\mathcal{A}, \Omega).$$

Pros and Cons of the Lifting Techniques

Pros:

- The lifting techniques is applicable to various alternative definitions of the Markovian JSR;
- The lifting techniques allows to investigate products of matrices defined by subshifts of finite type instead of Markov shifts;
- Potentially, the lifting techniques provides a possibility to apply the method of Barabanov norms to investigate Markovian products of matrices, however ***no works in this direction are known to me***.

Cons:

- All the matrices from $\mathcal{A}_\Omega$ are degenerate, and some of their products may turn to zero. This makes doubtful the application of the lifting techniques for studying the Markovian analogs of the LSR;
- It is unclear whether the lifting techniques may be applied to study infinite sets of matrices $\mathcal{A}$;

Further Results

- *Wang, Roohi, Dullerud & Viswanathan, 2014* — for matrix sequences generated by a Muller automaton;
- *Philippe, Essick, Dullerud & Jungers, 2015; Philippe & Jungers, 2015; Legat, Jungers & Parrilo, 2016* — for matrix sequences generated by general finite automata.

Frequency Constrained Matrix Products

Frequency...

Commonly used characteristics of the matrix products

$$A_{i_n} \cdots A_{i_2} A_{i_1}, \qquad A_{i_j} \in \mathcal{A} := \{A_1, A_2, \ldots, A_r\},$$

are the frequencies of occurrences of the indices $i \in 1, \ldots, r$ in the index sequence $\{i_n\}$. As a rule, given some $i \in 1, \ldots, r$, the frequency p_i is defined as the limit

$$p_i = \lim_{n \to \infty} p_{i,n}$$

of the relative frequencies (proportions)

$$p_{i,n} := \frac{\#\big(\, i_j \in \{i_1, i_2, \ldots, i_n\} : \; i_j = i \,\big)}{n}$$

of occurrences of the symbol i among the first n members of a sequence.

Frequency… (cont.)

The relative frequency $p_{i,n}$ for symbol i behaves as follows:

Deficiency of the Frequency Concept

- The definition of frequency ***is not enough informative*** since it does not answer the question of how often different symbols appear in intermediate, not tending to infinity, finite segments of a sequence.
- The definition of frequency becomes ***all the less satisfactory*** in situations when one should deal with not a single sequence but with an infinite collection of such sequences.
- The definition of frequency given above ***does not withstand transition to the limit*** with respect to different sequences which results in substantial theoretical and conceptual difficulties.

Deficiency of the Frequency Concept

To give "good" properties to determination of frequency one often needs:

- either to require some kind of uniformity of convergence of the relative frequencies $p_{i,n}$ to p_i
- or to treat appearance of the related symbols in a sequence as a realization of events generated by some random or deterministic ergodic system
- or something of this kind.

As a result, under such an approach ***one has to impose rather strong restrictions*** on the laws of forming the index sequences $\{i_n\}$ which are often difficult to verify of confirm in applications.

- The arising families of the index sequences and of the related matrix products can be rather attractive from the purely mathematical point of view but their description becomes less and less constructive.
- In applications, it leads to emergence of an essential conceptual gap or of some kind strained interpretation at use of the related objects and constructions.

What to do ?

Where to go ?

To be, or not to be ? †

Am I a trembling creature, or do I have the right ? ‡

. . .

† William Shakespeare. *Hamlet*

‡ Fyodor Dostoyevsky. *Crime and punishment*

Sequences with Constraints on the Sliding Block Frequencies

Let $p = (p_i, p_2, \ldots, p_r)$ be a set of positive numbers satisfying

$$p_1 + p_2 + \cdots + p_r = 1,$$

and let

$$p^- = (p_1^-, p_2^-, \ldots, p_r^-), \quad p^+ = (p_1^+, p_2^+, \ldots, p_r^+),$$

be sets of lower and upper bounds for p:

$$0 \le p_i^- < p_i < p_i^+ \le 1, \quad i = 1, 2, \ldots, r.$$

Sequences with Constraints on the Sliding Block Frequencies

Definition

Given a natural number ℓ, denote by $\mathcal{I}_\ell(p^\pm)$ the set of all infinite sequences $\{i_n\}_{n=0}^{\infty}$, $i_j \in \mathcal{I} := \{1, 2, \ldots, r\}$, for which the ***relative ℓ-block frequencies*** of occurrences of different symbols

$$p_{i,n}(\ell) := \frac{\#\left(i_j \in \{i_n, i_{n+1}, \ldots, i_{n+\ell-1}\} :\ i_j = i \right)}{\ell},$$

for each $i = 1, 2, \ldots, r$, satisfy

$$p_i^- \leq p_{i,n}(\ell) \leq p_i^+, \qquad \forall n.$$

$$i_0,\ i_1,\ \ldots\ ,\ \underbrace{i_n,\ i_{n+1},\ \ldots,\ i_{n+\ell-1}}_{\text{sliding } \ell\text{-block}},\ \ldots\ , i_k,\ i_{k+1},\ \ldots\ .$$

Example

Let $r = 3$, $\ell = 10$ and

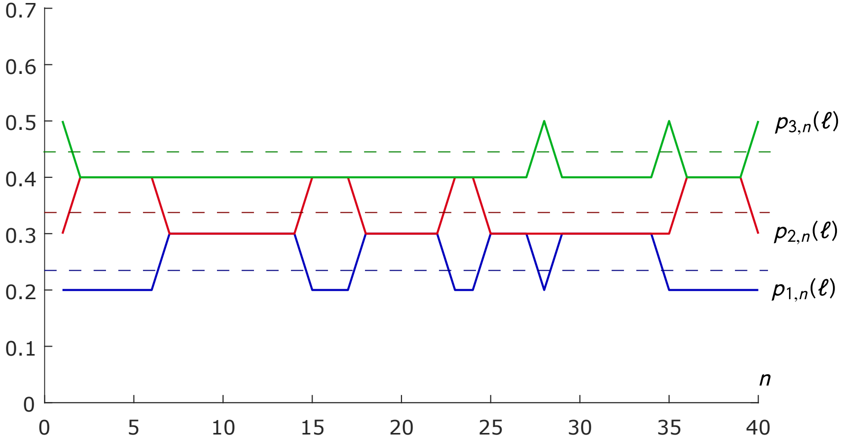
$$p = (0.23, 0.33, 0.44).$$

Define the sets of lower and upper bounds for p as follows:

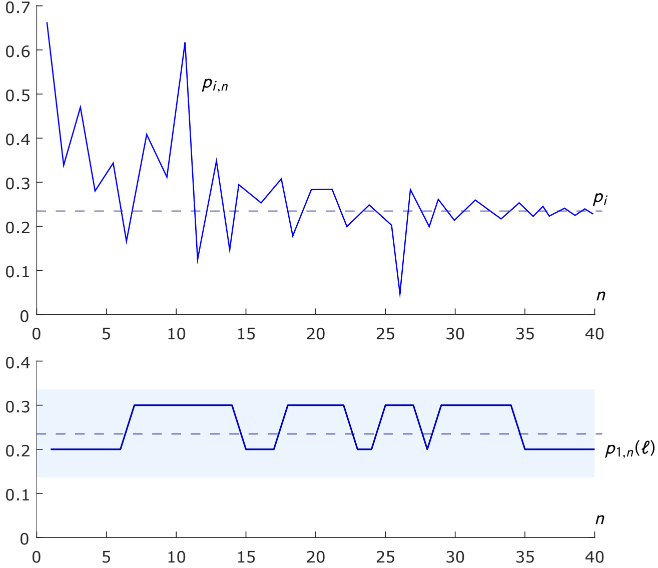
$$p^- = (0.13, 0.23, 0.34), \quad p^+ = (0.33, 0.43, 0.54).$$

Then the set $\mathcal{I}_\ell(p^\pm)$ contains the following sequences:

$$\boldsymbol{i}_1 = \{2, 1, 2, 3, 3, 3, 2, 3, 3, 1, 2, 1, 2, 3, 1, 3, 2, 3, 3, 3, 2, 1, 2, 2, 1, \ldots\},$$
$$\boldsymbol{i}_2 = \{3, 2, 2, 1, 3, 3, 3, 3, 2, 1, 2, 1, 2, 3, 3, 2, 3, 3, 2, 1, 2, 1, 3, 1, 3, \ldots\},$$
$$\boldsymbol{i}_3 = \{1, 1, 3, 3, 2, 2, 1, 2, 3, 3, 1, 3, 1, 3, 2, 2, 3, 2, 2, 3, 1, 3, 1, 3, 2, \ldots\}.$$


0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
5
10
15
20
25
30
35
40
n
$p_{3,n}(\ell)$
$p_{2,n}(\ell)$
$p_{1,n}(\ell)$

Difference Between Two Approaches



Main Properties

- The set $\mathcal{I}_\ell(p^\pm)$ ***is non-empty and "reach enough"*** if the "gaps" $p_i^+ - p_i^-$ are "not too small", e.g., $p_i^+ - p_i^- > \frac{2}{\ell}$, see (Kozyakin, 2014b) for details.
- In general, the frequencies of occurrences of the symbols $i = 1, 2, \ldots, r$ in the sequences from $\mathcal{I}_\ell(p^\pm)$ ***are not well defined***. The relative ℓ-block frequencies of the symbols $i = 1, 2, \ldots, r$ are "close" to the corresponding quantities p_i but, in general, they **may have no limits at infinity**.
- Given the sets $p^\pm$, the sequences from $\mathcal{I}_\ell(p^\pm)$ can be build constructively.

Theorem (Kozyakin, 2014b)

If $\mathcal{I}_\ell(p^\pm) \neq \varnothing$ then, for any sequence $\{i_n\}_{n=0}^{\infty} \in \mathcal{I}_\ell(p^\pm)$, transition between sequential sliding ℓ-blocks

$$\{i_n, i_{n+1}, \ldots, i_{n+\ell-1}\} \quad \Longrightarrow \quad \{i_{n+1}, i_{n+2}, \ldots, i_{n+\ell}\}$$

is a subshift of ℓ-type *or* ***an ℓ-step topological Markov chain****.*

Corollary

For the matrix products

$$A_{i_n} \cdots A_{i_2} A_{i_1}, \qquad A_{i_j} \in \mathcal{A} := \{A_1, A_2, \ldots, A_r\},$$

constrained to the index sequences $\{i_n\}_{n=0}^{\infty} \in \mathcal{I}_\ell(p^\pm)$ the JSR and GSR are well defined, and for them ***the Berger–Wang formula holds****.*

References

References

Barabanov, N. E. (1988).
On the Lyapunov exponent of discrete inclusions. I-III.
Automat. Remote Control, 49:152–157, 283–287, 558–565.

Berger, M. A. and Wang, Y. (1992).
Bounded semigroups of matrices.
Linear Algebra Appl., 166:21–27.

Blondel, V. D. and Nesterov, Y. (2005).
Computationally efficient approximations of the joint spectral radius.
SIAM J. Matrix Anal. Appl., 27(1):256–272 (electronic).

Bochi, J. and Morris, I. D. (2015).
Continuity properties of the lower spectral radius.
Proc. Lond. Math. Soc. (3), 110(2):477–509.

Bousch, T. and Mairesse, J. (2002).
Asymptotic height optimization for topical IFS, Tetris heaps, and the finiteness conjecture.
J. Amer. Math. Soc., 15(1):77–111 (electronic).

Chang, C.-T. and Blondel, V. D. (2013).
An experimental study of approximation algorithms for the joint spectral radius.
Numer. Algorithms, 64(1):181–202.

Chen, Q. and Zhou, X. (2000).
Characterization of joint spectral radius via trace.
Linear Algebra Appl., 315(1-3):175–188.

References (cont.)

Cicone, A. and Protasov, V. (2012).
Joint spectral radius computation.
MATLAB® Central.

Dai, X. (2011).
Optimal state points of the subadditive ergodic theorem.
Nonlinearity, 24(5):1565–1573.

Dai, X. (2012).
A Gel'fand-type spectral-radius formula and stability of linear constrained switching systems.
Linear Algebra Appl., 436(5):1099–1113.

Dai, X. (2014).
Robust periodic stability implies uniform exponential stability of Markovian jump linear systems and random linear ordinary differential equations.
J. Franklin Inst., 351(5):2910–2937.

Dai, X., Huang, Y., and Xiao, M. (2008).
Almost sure stability of discrete-time switched linear systems: a topological point of view.
SIAM J. Control Optim., 47(4):2137–2156.

Dai, X., Huang, Y., and Xiao, M. (2011a).
Periodically switched stability induces exponential stability of discrete-time linear switched systems in the sense of Markovian probabilities.
Automatica J. IFAC, 47(7):1512–1519.

References (cont.)

Dai, X., Huang, Y., and Xiao, M. (2011b).
Realization of joint spectral radius via ergodic theory.
Electron. Res. Announc. Math. Sci., 18:22–30.

Dai, X., Huang, Y., and Xiao, M. (2013).
Extremal ergodic measures and the finiteness property of matrix semigroups.
Proc. Amer. Math. Soc., 141(2):393–401.

Daubechies, I. and Lagarias, J. C. (1992).
Sets of matrices all infinite products of which converge.
Linear Algebra Appl., 161:227–263.

Elsner, L. (1995).
The generalized spectral-radius theorem: an analytic-geometric proof.
Linear Algebra Appl., 220:151–159.
Proceedings of the Workshop “Nonnegative Matrices, Applications and Generalizations” and the Eighth Haifa Matrix Theory Conference (Haifa, 1993).

Gripenberg, G. (1996).
Computing the joint spectral radius.
Linear Algebra Appl., 234:43–60.

Guglielmi, N. and Protasov, V. (2013).
Exact computation of joint spectral characteristics of linear operators.
Found. Comput. Math., 13(1):37–97.

References (cont.)

Guglielmi, N. and Zennaro, M. (2005).
Polytope norms and related algorithms for the computation of the joint spectral radius.
In *Proceedings of the 44th IEEE Conference on Decision and Control and European Control Conference 2005, Seville, Spain, December 12–15*, pages 3007–3012.

Gurvits, L. (1995).
Stability of discrete linear inclusion.
Linear Algebra Appl., 231:47–85.

Jungers, R. M. (2012).
On asymptotic properties of matrix semigroups with an invariant cone.
Linear Algebra Appl., 437(5):1205–1214.

Kozyakin, V. (2010).
Iterative building of Barabanov norms and computation of the joint spectral radius for matrix sets.
Discrete Contin. Dyn. Syst. Ser. B, 14(1):143–158.

Kozyakin, V. (2011).
A relaxation scheme for computation of the joint spectral radius of matrix sets.
J. Difference Equ. Appl., 17(2):185–201.

Kozyakin, V. (2013).
An annotated bibliography on convergence of matrix products and the theory of joint/generalized spectral radius.
Preprint, Institute for Information Transmission Problems, Moscow.

References (cont.)

Kozyakin, V. (2014a).
The Berger-Wang formula for the Markovian joint spectral radius.
Linear Algebra Appl., 448:315–328.

Kozyakin, V. (2014b).
Matrix products with constraints on the sliding block relative frequencies of different factors.
Linear Algebra Appl., 457:244–260.

Legat, B., Jungers, R. M., and Parrilo, P. A. (2016).
Generating unstable trajectories for switched systems via dual sum-of-squares techniques.
In *Proceedings of the 19th International Conference on Hybrid Systems: Computation and Control*, HSCC '16, pages 51–60, New York, NY, USA. ACM.

Maesumi, M. (1996).
An efficient lower bound for the generalized spectral radius of a set of matrices.
Linear Algebra Appl., 240:1–7.

Morris, I. D. (2010).
A rapidly-converging lower bound for the joint spectral radius via multiplicative ergodic theory.
Advances in Math., 225(6):3425–3445.

Morris, I. D. (2012).
The generalised Berger-Wang formula and the spectral radius of linear cocycles.
J. Funct. Anal., 262(3):811–824.

Morris, I. D. (2013).
Mather sets for sequences of matrices and applications to the study of joint spectral radii.
Proc. Lond. Math. Soc. (3), 107(1):121–150.

References (cont.)

Nesterov, Y. (2000).
Squared functional systems and optimization problems.
In *High performance optimization*, volume 33 of *Appl. Optim.*, pages 405–440. Kluwer Acad. Publ., Dordrecht.

Neumann, M. and Schneider, H. (1999).
The convergence of general products of matrices and the weak ergodicity of Markov chains.
Linear Algebra Appl., 287(1-3):307–314.
Special issue celebrating the 60th birthday of Ludwig Elsner.

Parrilo, P. A. (2000).
Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization.
PhD thesis, California Institute of Technology.

Parrilo, P. A. and Jadbabaie, A. (2007).
Approximation of the joint spectral radius of a set of matrices using sum of squares.
In *Hybrid systems: computation and control*, volume 4416 of *Lecture Notes in Comput. Sci.*, pages 444–458. Springer, Berlin.

Parrilo, P. A. and Jadbabaie, A. (2008).
Approximation of the joint spectral radius using sum of squares.
Linear Algebra Appl., 428(10):2385–2402.

Philippe, M., Essick, R., Dullerud, G., and Jungers, R. M. (2015).
Stability of discrete-time switching systems with constrained switching sequences.
ArXiv.org e-Print archive.

References (cont.)

Philippe, M. and Jungers, R. M. (2015).
Converse Lyapunov theorems for discrete-time linear switching systems with regular switching sequences.
In *Proceedings of the European Control Conference (ECC), 15-17 July 2015, Linz*. IEEE.

Protasov, V. Yu. (1996).
The joint spectral radius and invariant sets of linear operators.
Fundam. Prikl. Mat., 2(1):205–231.
in Russian.

Protasov, V. Y., Jungers, R. M., and Blondel, V. D. (2009/10).
Joint spectral characteristics of matrices: a conic programming approach.
SIAM J. Matrix Anal. Appl., 31(4):2146–2162.

Protasov, V. Yu. (2016).
Spectral simplex method.
Math. Program., 156(1-2, Ser. A):485–511.

Protasov, V. Y. and Jungers, R. M. (2012).
Convex optimization methods for computing the Lyapunov exponent of matrices.
ArXiv.org e-Print archive.

Rota, G.-C. and Strang, G. (1960).
A note on the joint spectral radius.
Nederl. Akad. Wetensch. Proc. Ser. A 63 = Indag. Math., 22:379–381.

References (cont.)

Shih, M.-H. (1999).
Simultaneous Schur stability.
Linear Algebra Appl., 287(1-3):323–336.
Special issue celebrating the 60th birthday of Ludwig Elsner.

Vankeerberghen, G., Hendrickx, J., Jungers, R., Chang, C. T., and Blondel, V. (2011).
The JSR Toolbox.
MATLAB® Central.

Vankeerberghen, G., Hendrickx, J., and Jungers, R. M. (2014).
JSR: A toolbox to compute the joint spectral radius.
In *Proceedings of the 17th International Conference on Hybrid Systems: Computation and Control*, HSCC'14, pages 151–156, New York, NY, USA. ACM.

Wang, Y., Roohi, N., Dullerud, G. E., and Viswanathan, M. (2014).
Stability of linear autonomous systems under regular switching sequences.
In *Proceedings of the 53d IEEE Conference on Decision and Control (CDC)*, pages 5445–5450.