Hourglass Alternative and constructivity of spectral characteristics of matrix products

VICTOR KOZYAKIN

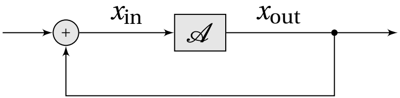
Kharkevich Institute
for Information Transmission Problems
Russian Academy of Sciences

Kotel'nikov Institute
of Radio-engineering and Electronics
Russian Academy of Sciences

Workshop on switching dynamics & verification
Amphithéâtre Darboux, Institut Henry Poincaré (IHP), Paris, France,
January 28–29, 2016.

Introduction

Main point of interest: stability/stabilizability of a discrete-time system



described by a linear (*switching*) equation

$$x(n+1) = A(n)x(n), \quad n = 0, 1, \ldots,$$

where

$$A(n) \in \mathscr{A} = \{A_1, A_2, \ldots, A_r\}, \quad A_i \in \mathbb{R}^{d\times d},$$
$$x(n) \in \mathbb{R}^d.$$

General problem

This problem is a special case of the more general problem:

When the matrix products $A_{i_n}\cdots A_{i_2}A_{i_1}$ with $i_n\in\{1,\ldots,r\}$ converge under different assumptions on the switching sequences $\{i_n\}$?

- "parallel" vs "sequential" computational algorithms: e.g., Gauss-Seidel vs Jacobi method;
- distributed computations;
- "asynchronous" vs "synchronous" mode of data exchange in the control theory and data transmission (large-scale networks);
- smoothness problems for Daubeshies wavelets (computational mathematics);
- one-dimensional discrete Schrödinger equations with quasiperiodic potentials (theory of quasicrystalls, physics);
- linear or affine iterated function systems (theory of fractals);
- Hopfield-Tank neural networks (biology, mathematics);
- "triangular arbitrage" in the models of market economics;
- etc.

Joint and Lower Spectral Radii

Given a set of $(d \times d)$-matrices $\mathscr{A}$ and a norm $\|\cdot\|$ on $\mathbb{R}^d$,

$$\rho(\mathscr{A}) = \lim_{n\to\infty} \sup\left\{\|A_{i_n}\cdots A_{i_1}\|^{1/n} :~ A_{i_j} \in \mathscr{A}\right\}$$

is called the *joint spectral radius (JSR)* of $\mathscr{A}$ (Rota & Strang, 1960), whereas

$$\check{\rho}(\mathscr{A}) = \lim_{n\to\infty} \inf\left\{\|A_{i_n}\cdots A_{i_1}\|^{1/n} :~ A_{i_j} \in \mathscr{A}\right\}$$

is called the *lower spectral radius (LSR)* of $\mathscr{A}$ (Gurvits, 1995).

Remark

- $\rho(\mathscr{A})$ and $\check{\rho}(\mathscr{A})$ are well defined and independent on the norm $\|\cdot\|$;
- $\|\cdot\|$ in the definitions of JSR and LSR may be replaced by the spectral radius $\rho(\cdot)$ of a matrix, see Berger & Wang, 1992 for $\rho(\mathscr{A})$ and Gurvits, 1995; Theys, 2005; Czornik, 2005 for $\check{\rho}(\mathscr{A})$.

Another Formulae for JSR

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

Stability vs Stabilizability

Difference between the joint and lower spectral radii:

The inequality $\rho(\mathscr{A}) < 1$ characterizes the ***Schur stability*** of $\mathscr{A}$:

$$\rho(\mathscr{A}) < 1 \quad\Longrightarrow\quad \forall\{i_n\}: \quad \|A_{i_n}\cdots A_{i_2}A_{i_1}\| \to 0.$$

The inequality $\check{\rho}(\mathscr{A}) < 1$ characterizes the ***Schur stabilizability*** of $\mathscr{A}$:

$$\check{\rho}(\mathscr{A}) < 1 \quad\Longrightarrow\quad \exists\{i_n\}: \quad \|A_{i_n}\cdots A_{i_2}A_{i_1}\| \to 0.$$

JSR vs LSR

- The LSR has ‘less stable’ continuity properties than the JSR, see Bousch & Mairesse, 2002;

- Until recently, ‘good’ properties for 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.

  Their investigation is based on the concepts of *dominated splitting* and *k-multicones* from the theory of hyperbolic linear cocycles. In particular, they gave a **sufficient condition for the Lipschitz continuity of the LSR**

First Problems

Inequalities

$$\rho(\mathscr{A}) < 1, \qquad \check{\rho}(\mathscr{A}) < 1$$

might seem to give an exhaustive answer to the questions on stability or stabilizability of a switching system.

Theoretically:

this is indeed the case.

In practice:

- the computation of $\rho(\mathscr{A})$ and $\check{\rho}(\mathscr{A})$ is generally **impossible in a closed formula form** $\Longrightarrow$ need in approximate computational methods;
- there are **no a priory estimates** for the rate of convergence of the related limits in the definitions of $\rho(\mathscr{A})$ and $\check{\rho}(\mathscr{A})$;
- the required **amount of computations rapidly increases** in n and dimension of a system.

First Problems (cont.)

The following problems of stability and stabilizability of linear switching systems are not new per se, but are remaining to be relevant.

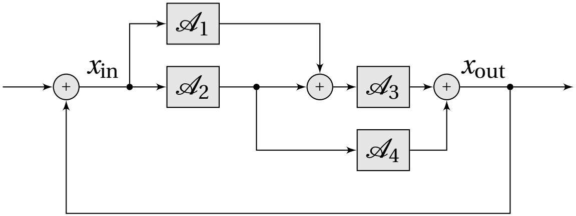
Problem

How to describe the classes of switching systems (classes of matrix sets $\mathscr{A}$), for which the JSR $\rho(\mathscr{A})$ could be constructively calculated?

Problem

How to describe the classes of switching systems (classes of matrix sets $\mathscr{A}$), for which the LSR $\check{\rho}(\mathscr{A})$ could be constructively calculated?

It is of crucial importance that in the control theory, in general, **systems are composed not of a single block but of a number of interconnected blocks**, e.g.



When these blocks are linear and functioning asynchronously, each of them is described by the equation

$$x_{\text{out}}(n+1) = A_i(n)x_{\text{in}}(n), \qquad x_{\text{in}}(\cdot) \in \mathbb{R}^{N_i},\ x_{\text{out}}(\cdot) \in \mathbb{R}^{M_i},\ \ n = 0, 1, \ldots,$$

where the matrices $A_i(n)$, for each n, may arbitrarily take values from some set $\mathscr{A}_i$ of $(N_i \times M_i)$-matrices, where $i = 1, 2, \ldots, Q$ and Q is the total amount of blocks.

Another Problem (cont.)

Question

What can be said about stability or stabilizability of a system, whose blocks may be connected in parallel or in series, or in a more complicated way, represented by some directed graph with blocks placed on its edges?

Disappointing Remark:

Under such a connection of blocks, the classes of matrices describing the transient processes of a system as a whole became very complicated and their properties are practically not investigated.

So, the following problem is also urgent:

Problem

How to describe the switching systems for which the question about stability or stabilizability can be constructively answered not only for an isolated switching blocks but also for any series-parallel connection of such blocks?

Constructive computability of spectral characteristics

Finiteness Conjecture

The possibility of 'explicit' calculation of the spectral characteristics of sets of matrices is conventionally associated with the validity of the *finiteness conjecture* (Lagarias & Wang, 1995) according to which *the limit in the formulas*

$$\rho(\mathscr{A}) = \lim_{n\to\infty} \sup\left\{\|A_{i_n}\cdots A_{i_1}\|^{1/n}:~ A_{i_j}\in\mathscr{A}\right\},$$

$$\check{\rho}(\mathscr{A}) = \lim_{n\to\infty} \inf\left\{\|A_{i_n}\cdots A_{i_1}\|^{1/n}:~ A_{i_j}\in\mathscr{A}\right\}$$

is attained at some finite value of n.

This finiteness conjecture was disproved

- for JSR: Bousch & Mairesse, 2002. The 'explicit' counterexamples to the finiteness conjecture was built by Hare, Morris, Sidorov & Theys, 2011; Morris & Sidorov, 2013; Jenkinson & Pollicott, 2015.
- for LSR: Bousch & Mairesse, 2002; Czornik & Jurgaś, 2007.

Finiteness Conjecture (cont.)

Despite the finiteness conjecture is false, attempts to discover new classes of matrices for which it still occurs continues.

Should be borne in mind:

The validity of the finiteness conjecture for some class of matrices provides **only a theoretical possibility** to 'explicitly' calculate the related spectral characteristics, because in practice calculation of the spectral radii $\rho(A_n \cdots A_1)$ for all possible sets of matrices $A_1, \ldots, A_n \in \mathscr{A}$ may require too much computing resources, even for relatively small values of n.

$$\Downarrow$$

From the practical point of view, the most interesting are the cases when the finiteness conjecture holds for small values of n.

Finiteness Conjecture (cont.)

The Finiteness Conjecture is known to be valid in the following cases:

- $\mathscr{A}$ is a set of commuting matrices;
- $\mathscr{A}$ is a set of upper or lower triangular matrices
- $\mathscr{A}$ is a set of isometries in some norm up to a scalar factor (that is, $\|Ax\| \equiv \lambda_A \|x\|$ for some λ_A).
- $\mathscr{A}$ is a 'symmetric' bounded set of matrices: together with each matrix $\mathscr{A}$ contains also the (complex) conjugate matrix (Plischke & Wirth, 2008). This class includes all the sets of self-adjoint matrices.
- $\mathscr{A}$ is a set of the so-called non-negative matrices *with independent row uncertainty* (Blondel & Nesterov, 2009).
- $\mathscr{A}$ is a pair of 2×2 binary matrices, i.e. matrices with the elements $\{0,1\}$ (Jungers & Blondel, 2008).
- $\mathscr{A}$ is a pair of 2×2 sign-matrices, i.e. matrices with the elements $\{-1,0,1\}$ (Cicone, Guglielmi, Serra-Capizzano & Zennaro, 2010).
- $\mathscr{A}$ is a bounded family of matrices, whose matrices, except perhaps one, have rank 1 (Morris, 2011; Dai, Huang, Liu & Xiao, 2012; Liu & Xiao, 2012; Liu & Xiao, 2013; Wang & Wen, 2013).

Theorem (Blondel & Nesterov, 2009)

Both $\rho(\mathscr{A})$ and $\check{\rho}(\mathscr{A})$ can be constructively calculated provided that $\mathscr{A}$ is a ***set of non-negative matrices with independent row uncertainty***.

Definition (Blondel & Nesterov, 2009)

A set of $N \times M$-matrices $\mathscr{A}$ is called a *set with independent row uncertainty*, or an *IRU-set*, if it consists of all the matrices

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1M} \\ a_{21} & a_{22} & \cdots & a_{2M} \\ \dots & \dots & \dots & \dots \\ a_{N1} & a_{N2} & \cdots & a_{NM} \end{pmatrix},$$

each row $a_i = (a_{i1}, a_{i2}, \ldots, a_{iM})$ of which belongs to some set of M-rows $\mathscr{A}^{(i)}$, $i = 1, 2, \ldots, N$.

Example

Let the sets of rows $\mathscr{A}^{(1)}$ and $\mathscr{A}^{(2)}$ be as follows:

$$\mathscr{A}^{(1)} = \{(a,b),\ (c,d)\},\quad \mathscr{A}^{(2)} = \{(\alpha,\beta),\ (\gamma,\delta),\ (\mu,\nu)\}.$$

Then the IRU-set $\mathscr{A}$ consists of the following matrices:

$$A_{11} = \begin{pmatrix} a & b \\ \alpha & \beta \end{pmatrix},\quad A_{12} = \begin{pmatrix} a & b \\ \gamma & \delta \end{pmatrix},\quad A_{13} = \begin{pmatrix} a & b \\ \mu & \nu \end{pmatrix},$$

$$A_{21} = \begin{pmatrix} c & d \\ \alpha & \beta \end{pmatrix},\quad A_{22} = \begin{pmatrix} c & d \\ \gamma & \delta \end{pmatrix},\quad A_{23} = \begin{pmatrix} c & d \\ \mu & \nu \end{pmatrix}.$$

Example

Let

$$\mathscr{A}^{(1)} = \{(a_{11}, a_{12}),\ (1,0)\}, \quad \mathscr{A}^{(2)} = \{(a_{21}, a_{22}),\ (0,1)\}.$$

Then the IRU-set $\mathscr{A}$ consists of the following matrices:

$$A_{11} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},\ A_{12} = \begin{pmatrix} a_{11} & a_{12} \\ 0 & 1 \end{pmatrix},\ A_{21} = \begin{pmatrix} 1 & 0 \\ a_{21} & a_{22} \end{pmatrix},\ A_{22} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Matrices of such a kind are known long ago in the computational mathematics and control theory:

- matrices A_{12}, A_{21} are used in place of A_{11} during transition from ‘parallel’ to ‘sequential’ computational algorithms: e.g., from the Jacobi method to the Gauss-Seidel one;
- matrices A_{ij} arise in the control theory in description of ‘data loss’ information exchange.

Finiteness Theorem (Blondel & Nesterov, 2009; Nesterov & Protasov, 2013)

If an IRU-set of non-negative matrices $\mathscr{A}$ is compact then

$$\rho(\mathscr{A}) = \max_{A\in\mathscr{A}} \rho(A), \qquad \check{\rho}(\mathscr{A}) = \min_{A\in\mathscr{A}} \rho(A).$$

Remark

For IRU-sets of arbitrary matrices, the Blondel-Nesterov-Protasov theorem is not valid.

Hourglass Alternative

The original proof of the Finiteness Theorem is quite cumbersome, so outline the idea of alternative proof (Kozyakin, 2016).

Main observation: for IRU-sets of non-negative matrices the following assertion holds:

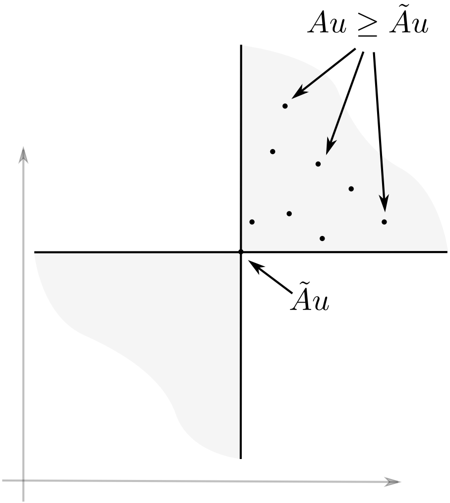
Hourglass Alternative

Given a matrix $\tilde{A} \in \mathscr{A}$ and a vector $u > 0$

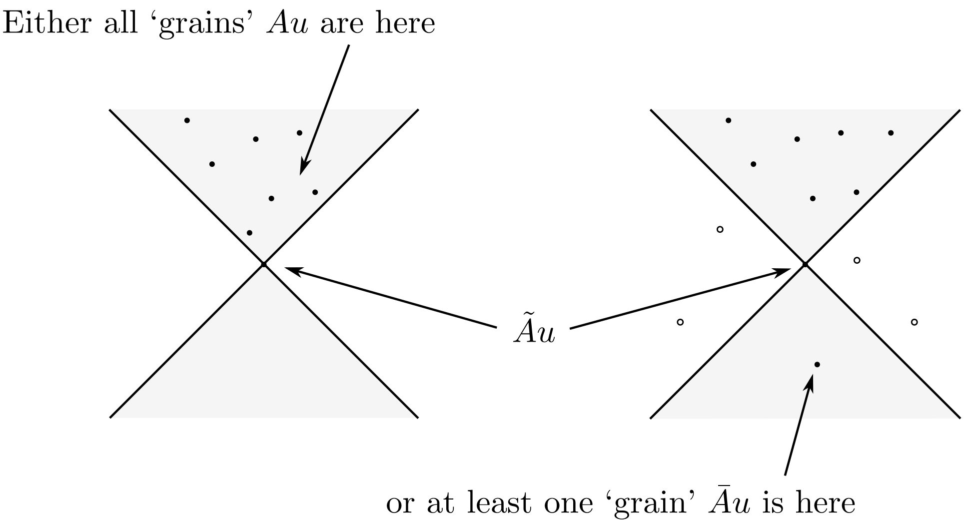
$$\Downarrow$$

H1: either $Au \ge \tilde{A}u$ for all $A \in \mathscr{A}$ or $\exists\, \bar{A} \in \mathscr{A}$: $\bar{A}u \le \tilde{A}u$ and $\bar{A}u \ne \tilde{A}u$;

H2: either $Au \le \tilde{A}u$ for all $A \in \mathscr{A}$ or $\exists\, \bar{A} \in \mathscr{A}$: $\bar{A}u \ge \tilde{A}u$ and $\bar{A}u \ne \tilde{A}u$.



Rotate this Figure 45° counterclockwise!


Either all ‘grains’ Au are here
$\tilde{A}u$
or at least one ‘grain’ $\bar{A}u$ is here

Idea of Proof of the Blondel-Nesterov-Protasov Theorem

Let $\tilde{A} \in \mathscr{A}$ be such that $\rho(\tilde{A}) = \max_{A \in \mathscr{A}} \rho(A)$ and $u > 0$ be the leading eigenvalue of $\tilde{A}$.

Hourglass Alternative

$$\Downarrow$$

$$Au \le \tilde{A}u \qquad \forall A \in \mathscr{A}$$

$$\Downarrow$$

$$A_{i_n} \cdots A_{i_1} u \le \tilde{A}^n u \qquad \forall A_{i_j} \in \mathscr{A}$$

$$\Downarrow$$

$$\rho(A_{i_n} \cdots A_{i_1}) \le \rho^n(\tilde{A}) \qquad \forall A_{i_j} \in \mathscr{A}$$

$$\Downarrow$$

$$\rho(\mathscr{A}) \le \max_{A \in \mathscr{A}} \rho(A) \qquad (\le \rho(\mathscr{A}))$$

The Hourglass Alternative is **the only property which was used in the proof** of the Blondel-Nesterov-Protasov Finiteness theorem! So,

Let us axiomatize this property!

Definition (Kozyakin, 2016)

A set of positive matrices $\mathscr{A}$ is called an *$\mathscr{H}$-set*, if it satisfies the Hourglass Alternative.

Example

- any IRU-set of positive matrices is an $\mathscr{H}$-set;
- any set of positive matrices $\mathscr{A} = \{A_1, A_2, \ldots, A_n\}$ satisfying $A_1 \le A_2 \le \cdots \le A_n$ (called *linearly ordered* set) is an $\mathscr{H}$-set.

Not every set of positive matrices is an $\mathscr{H}$-set.

Recall the Minkowski operations of addition and multiplication for sets of matrices:

$$\mathscr{A} + \mathscr{B} := \{A + B : A \in \mathscr{A},\ B \in \mathscr{B}\},$$

$$\mathscr{A}\mathscr{B} := \{AB : A \in \mathscr{A},\ B \in \mathscr{B}\},$$

$$t\mathscr{A} = \mathscr{A}t := \{tA : t \in \mathbb{R},\ A \in \mathscr{A}\}$$

Remark on the Operations of Minkowski

The addition of sets of matrices corresponds to the **parallel coupling** of independently operating asynchronous controllers functioning independently.

The multiplication corresponds to the **serial coupling** of asynchronous controllers.

Denote the totality of all $\mathscr{H}$-sets of $(N \times M)$-matrices by $\mathscr{H}(N,M)$.

Theorem (Kozyakin, 2016)

The following is true:

(i) $\mathscr{A} + \mathscr{B} \in \mathscr{H}(N,M)$, *if* $\mathscr{A}, \mathscr{B} \in \mathscr{H}(N,M)$;

(ii) $\mathscr{A}\mathscr{B} \in \mathscr{H}(N,Q)$, *if* $\mathscr{A} \in \mathscr{H}(N,M)$ *and* $\mathscr{B} \in \mathscr{H}(M,Q)$;

(iii) $t\mathscr{A} = \mathscr{A}t \in \mathscr{H}(N,M)$, *if* $t > 0$ *and* $\mathscr{A} \in \mathscr{H}(N,M)$.

The totality $\mathscr{H}(N,N)$ is endowed with additive and multiplicative group operations, but itself is not a group, neither additive nor multiplicative.

After adding the zero additive element $\{0\}$ and the identity multiplicative element $\{I\}$ to $\mathscr{H}(N,N)$, the resulting totality $\mathscr{H}(N,N) \cup \{0\} \cup \{I\}$ becomes a semiring in the sense of Golan, 1999.

Remark

By the above theorem any finite sum of any finite products of sets of matrices from $\mathscr{H}(N,N)$ is again a set of matrices from $\mathscr{H}(N,N)$. Moreover, for any integers $n,d\ge 1$, all the polynomial sets of matrices

$$P(\mathscr{A}_1,\mathscr{A}_1,\dots,\mathscr{A}_n)=\sum_{k=1}^{d}\sum_{i_1,i_2,\dots,i_k\in\{1,2,\dots,n\}} p_{i_1,i_2,\dots,i_k}\mathscr{A}_{i_1}\mathscr{A}_{i_2}\cdots\mathscr{A}_{i_k},$$

where $\mathscr{A}_1,\mathscr{A}_1,\dots,\mathscr{A}_n\in\mathscr{H}(N,N)$ and the scalar coefficients $p_{i_1,i_2,\dots,i_k}$ are positive, belong to the set $\mathscr{H}(N,N)$.

Theorem (Kozyakin, 2016)

Let $\mathscr{A}\in\overline{\mathscr{H}}(N,N)$. Then

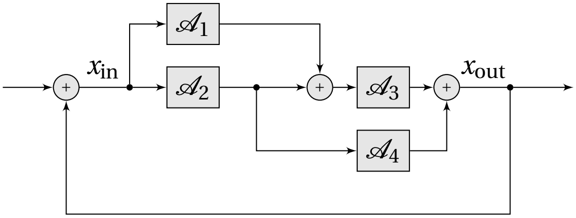
$$\rho(\mathscr{A})=\max_{A\in\mathscr{A}}\rho(A),\qquad \check{\rho}(\mathscr{A})=\min_{A\in\mathscr{A}}\rho(A).$$

What does it imply for the control theory?

Theorem

Given a system formed by a series-parallel connection of blocks corresponding to some $\mathscr{H}$-sets of non-negative matrices $\mathscr{A}_i$, $i = 1, 2, \ldots, Q$.

Then the question of stability (stabilizability) of such a system can be constructively resolved by finding a matrix at which $\max_{A \in \mathscr{A}} \rho(A)$ is attained, where $\mathscr{A}$ is the Minkowski polynomial sum of the matrix sets $\mathscr{A}_i$, $i = 1, 2, \ldots, Q$, corresponding to the structure of coupling of the related blocks.



Questions

- Any other examples of $\mathscr{H}$-sets of matrices?
- Is it possible to extend this approach to non-positive matrices?
- What can be said about control systems with non-directed coupling of blocks?
- etc.

Individual Trajectories

One More Problem

Both, the JSR and the LSR of a matrix set, describe the limiting behavior of the 'multiplicatively averaged' norms of the matrix products, $\|A_{i_n}\cdots A_{i_1}\|^{1/n}$. That is,

they characterize the stability or stabilizability of a system 'as a whole'.

Often there arise the problem to find, **for a given *x***, a sequence of matrices that would ensure the fastest 'increase or decrease' of the quantities

$$\nu(A_{i_n}\cdots A_{i_1}x),$$

where $\nu(\cdot)$ is a numerical function.

Examples of the function $\nu(\cdot)$ are the norms

$$\|x\|_1=\sum_i|x_i|,\quad \|x\|_2=\sqrt{\sum_i|x_i|^2},\quad \|x\|_\infty=\max_i|x_i|.$$

One More Problem (cont.)

If $\mathscr{A}$ is a finite set consisting of K elements then to find the value of

$$\max_{A_{i_j}\in\mathscr{A}} \nu(A_{i_n}\cdots A_{i_1}x) \qquad (*)$$

one need, in general, to compute K^n times the values of the function $\nu(\cdot)$.

Problem

How to describe the classes of switching systems (the classes of matrix sets $\mathscr{A}$), for which the number of computations of $\nu(\cdot)$ needed to calculate the quantity $(*)$ would be less than K^n?

It is desirable that the required number of computations would be of order Kn.

A similar problem on minimization of $\nu(A_{i_n}\cdots A_{i_1}x)$ can also be posed.

One-step Maximization

First consider the problem of finding

$$\max_{A\in\mathscr{A}} \nu(Ax),$$

where $\mathscr{A}$ is assumed to be compact.

By Assertion H2 of the Hourglass Alternative, for any matrix $\tilde{A}\in\mathscr{A}$, either $Ax\le\tilde{A}x$ for all $A\in\mathscr{A}$ or there exists a matrix $\bar{A}\in\mathscr{A}$ such that $\bar{A}x\ge\tilde{A}x$ and $\bar{A}x\ne\tilde{A}x$.

This, together with the compactness of the set $\mathscr{A}$, implies the existence of a matrix $A_x^{(max)}\in\mathscr{A}$ such that,

$$\boxed{Ax\le A_x^{(max)}x,\qquad \forall\ A\in\mathscr{A}.}$$

One-step Maximization (cont.)

Theorem

Let $\mathscr{A}$ be a compact $\mathscr{H}$-set of non-negative $(N\times N)$-matrices, $\nu(\cdot)$ be a coordinate-wise monotone function, and $x\in\mathbb{R}^N$, $x\ge 0$, be a vector.

(i) *Then*

$$\max_{A\in\mathscr{A}} \nu(Ax) = \nu(A_x^{(max)}x).$$

(ii) *Let, additionally, the function $\nu(\cdot)$ be strictly coordinate-wise monotone. If*

$$\max_{A\in\mathscr{A}} \nu(Ax) = \nu(\tilde{A}x)$$

then

$$\tilde{A}x = A_x^{(max)}x.$$

Multi-step Maximization

We turn now to the question of determining the quantity $\nu(A_{i_n}\cdots A_{i_1}x)$ for some $n>1$ and $x\in\mathbb{R}^N$, $x\geq 0$. With this aim in view, let us construct sequentially the matrices $A_i^{(max)}$, $i=1,2,\ldots,n$, as follows:

- the matrix $A_1^{(max)}$ is constructed in the same way as was done in the previous section: $A_1^{(max)}=A_{x_0}^{(max)}$;
- if the matrices $A_i^{(max)}$, $i=1,2,\ldots,k$, have already constructed then the matrix $A_{k+1}^{(max)}$, depending on the vector

$$x_k=A_k^{(max)}\cdots A_1^{(max)}x,$$

is constructed to maximize the function

$$\nu(AA_k^{(max)}\cdots A_1^{(max)}x)=\nu(Ax_k)$$

over all $A\in\mathscr{A}$ in the same manner as was done in the previous section. So, the matrix $A_{k+1}^{(max)}$ is defined by the equality $A_{k+1}^{(max)}=A_{x_k}^{(max)}$.

Multi-step Maximization (cont.)

Theorem

Let $\mathscr{A}$ be a compact $\mathscr{H}$-set of non-negative $(N\times N)$-matrices, $\nu(\cdot)$ be a coordinate-wise monotone function, and $x\in\mathbb{R}^N$, $x\ge 0$, be a vector.

(i) *Then*

$$\max_{A_n,\ldots,A_1\in\mathscr{A}} \nu(A_n\cdots A_1x)=\nu(A_n^{(max)}\cdots A_1^{(max)}x).$$

(ii) *Let, additionally, the set $\mathscr{A}$ consist of positive matrices and the function $\nu(\cdot)$ be strictly coordinate-wise monotone. If*

$$\max_{A_n,\ldots,A_1\in\mathscr{A}} \nu(A_n\cdots A_1x)=\nu(\tilde{A}_n\cdots\tilde{A}_1x)$$

then

$$\tilde{A}_i\cdots\tilde{A}_1x=A_i^{(max)}\cdots A_i^{(max)}x,\qquad i=1,2,\ldots,n.$$

Minimax Theorem

Minimax Theorem

Let $\mathscr{A}\in\overline{\mathscr{H}}(N,M)$ and $\mathscr{B}\in\overline{\mathscr{H}}(M,N)$. Then

$$\min_{A\in\mathscr{A}}\max_{B\in\mathscr{B}}\rho(AB)=\max_{B\in\mathscr{B}}\min_{A\in\mathscr{A}}\rho(AB).$$

Asarin, Cervelle, Degorre, Dima, Horn & Kozyakin, 2015 used a restricted form of this theorem to investigate the so-called *matrix multiplication games* (to be presented at STACS 2016, Orléans, France, February 17-20).

Remark

In the Minimax Theorem, $\mathscr{A}$ and $\mathscr{B}$ may be replaced by any compact subsets of $\operatorname{conv}(\mathscr{A})$ and $\operatorname{conv}(\mathscr{B})$, respectively.

Minimax Theorem: Difficulty of Proof

1. The vast majority of proofs of the minimax theorems heavily employ some kind of **convexity in one of the arguments** of the related function **and concavity in the other** (see, e.g., survey Simons, 1995).
2. We were not able to find suitable analogs of convexity or concavity of the function $\rho(AB)$ with respect to the matrix variables A and B.
3. In our context, due to the identity

$$\rho(AB) \equiv \rho(BA),$$

the role of the matrices A and B is in a sense equivalent. Therefore, any kind of 'convexity' of the function $\rho(AB)$ with respect, say, to the variable A would have to involve its 'concavity' with respect to the same variable, which casts doubt on the applicability of the 'convex-concave' arguments in the proof of the Minimax Theorem.

The work was carried out at the Kotel'nikov Institute of Radio-engineering and Electronics, Russian Academy of Sciences, and was funded by the Russian Science Foundation, Project No. 16-11-00063.

References

Asarin, E., Cervelle, J., Degorre, A., Dima, C., Horn, F., and Kozyakin, V. (2015).
Entropy games and matrix multiplication games.
ArXiv.org e-Print archive.

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).

Blondel, V. D. and Nesterov, Y. (2009).
Polynomial-time computation of the joint spectral radius for some sets of nonnegative matrices.
SIAM J. Matrix Anal. Appl., 31(3):865–876.

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

References (cont.)

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).

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

Cicone, A., Guglielmi, N., Serra-Capizzano, S., and Zennaro, M. (2010).
Finiteness property of pairs of 2×2 sign-matrices via real extremal polytope norms.
Linear Algebra Appl., 432(2-3):796–816.

Czornik, A. (2005).
On the generalized spectral subradius.
Linear Algebra Appl., 407:242–248.

Czornik, A. and Jurgaś, P. (2007).
Falseness of the finiteness property of the spectral subradius.
Int. J. Appl. Math. Comput. Sci., 17(2):173–178.

Dai, X., Huang, Y., Liu, J., and Xiao, M. (2012).
The finite-step realizability of the joint spectral radius of a pair of $d\times d$ matrices one of which being rank-one.
Linear Algebra Appl., 437(7):1548–1561.

References (cont.)

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).

Golan, J. S. (1999).
Semirings and their applications.
Kluwer Academic Publishers, Dordrecht.

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

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

Hare, K. G., Morris, I. D., Sidorov, N., and Theys, J. (2011).
An explicit counterexample to the Lagarias-Wang finiteness conjecture.
Adv. Math., 226(6):4667–4701.

Jenkinson, O. and Pollicott, M. (2015).
Joint spectral radius, Sturmian measures, and the finiteness conjecture.
ArXiv.org e-Print archive.

References (cont.)

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

Jungers, R. M. and Blondel, V. D. (2008).
On the finiteness property for rational matrices.
Linear Algebra Appl., 428(10):2283–2295.

Kozyakin, V. (2016).
Hourglass alternative and the finiteness conjecture for the spectral characteristics of sets of non-negative matrices.
Linear Algebra Appl., 489:167–185.

Lagarias, J. C. and Wang, Y. (1995).
The finiteness conjecture for the generalized spectral radius of a set of matrices.
Linear Algebra Appl., 214:17–42.

Liu, J. and Xiao, M. (2012).
Computation of joint spectral radius for network model associated with rank-one matrix set.
In *Neural Information Processing. Proceedings of the 19th International Conference, ICONIP 2012, Doha, Qatar, November 12-15, 2012, Part III*, volume 7665 of *Lecture Notes in Computer Science*, pages 356–363. Springer Berlin Heidelberg.

Liu, J. and Xiao, M. (2013).
Rank-one characterization of joint spectral radius of finite matrix family.
Linear Algebra Appl., 438(8):3258–3277.

References (cont.)

Morris, I. and Sidorov, N. (2013).
On a Devil's staircase associated to the joint spectral radii of a family of pairs of matrices.
J. Eur. Math. Soc. (JEMS), 15(5):1747–1782.

Morris, I. D. (2011).
Rank one matrices do not contribute to the failure of the finiteness property.
ArXiv.org e-Print archive.

Nesterov, Y. and Protasov, V. Y. (2013).
Optimizing the spectral radius.
SIAM J. Matrix Anal. Appl., 34(3):999–1013.

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

Plischke, E. and Wirth, F. (2008).
Duality results for the joint spectral radius and transient behavior.
Linear Algebra Appl., 428(10):2368–2384.

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

References (cont.)

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.

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.

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

Simons, S. (1995).
Minimax theorems and their proofs.
In *Minimax and applications*, volume 4 of *Nonconvex Optim. Appl.*, pages 1–23. Kluwer Acad. Publ., Dordrecht.

Theys, J. (2005).
Joint Spectral Radius: Theory and Approximations.
PhD thesis, Faculté des sciences appliquées, Département d'ingénierie mathématique, Center for Systems Engineering and Applied Mechanics, Université Catholique de Louvain.

References (cont.)

Wang, S. and Wen, J. (2013).
The finiteness conjecture for the joint spectral radius of a pair of matrices.
In *Proceedings of the 9th International Conference on Computational Intelligence and Security (CIS), 2013, Emeishan, China, December 14–15*, pages 798–802.