Vladimir Protasov

L'Aquila (Italy), MSU, HSE (Russia)

# Four stories on the stability of linear systems

# Story 3

## Positive linear systems.

How to find the closest stable/unstable system?

A linear dynamical system with discrete time:

 $x_{k+1} = A x_k, \quad k \in \mathbb{Z}_+ \quad x_0 \text{ is given}$ 

A system is stable if all trajectories tend to zero (Schur stability)



How to find the closest stable matrix to A? How far A is to the set of stable matrices? How to find the closest unstable matrix to A? How far A is to the set of unstable matrices?

$$||X - A|| \to \min$$
  

$$\rho(X) = 1$$

This is equivalent to the optimizing of the spectral radius of a matrix over a matrix ball

Let A be a d x d matrix,  $\rho(A)$  be its spectral radius.

If  $\lambda_1, ..., \lambda_d$  are eigenvalues of *A* and  $|\lambda_1| \ge ... \ge |\lambda_d|$ , then  $\rho(A) = |\lambda_1|$ .

Theorem 1. (Perron (1906), Frobenius (1913)). If  $A \ge 0$ , then the spectral radius is attained at a real positive eigenvalue  $\rho(A) = \lambda_1 \ge 0$ .

There is an eigenvector  $v \ge 0$  such that  $Av = \lambda_1 v$ .

If A > 0, then the largest by modulo eigenvalue is unique and simple, and so is the corresponding eigenvector.

$$\rho(A) = \lambda_1$$

We call  $\lambda_{max} = \lambda_1 = \rho(A)$  the leading eigenvalue and v the leading eigenvector.

## **Optimizing the spectral radius of a matrix**

The general problem:

$$\left( \begin{array}{c} \rho(A) \to \max / \min \\ A \in M \end{array} \right)$$

M is a set of marices.

The nonnegative case:

$$\begin{cases} \lambda_{\max}(A) \to \max/\min\\ A \in M \end{cases}$$

M is a set of nonnegative marices.

**These problems are both notoriously hard** (even if the set M is convex).

#### **Reasons:**

The spectral radius is neither convex nor concave in matrices

The spectral radius is non-Lipschitz, if the leading eigenvalue is multiple.

**Example 1.** For the set  $M = [A_1, A_2] = co\{A_1, A_2\},$ 

$$A_{1} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0.1 \\ 0 & 0 & 0.1 & 0 \end{pmatrix} \quad ; \quad A_{2} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.1 \\ 0 & 0 & 0.1 & 0.5 \end{pmatrix}$$

We have 
$$A = (1-x)A_1 + x A_2$$
,  $x \in [0,1]$ . "Matrix segment"



We consider the closest stable nonnegative matrix in two polyhedral forms in  $\mathbb{R}^d$ :

The 
$$L_{\infty}$$
 - norm:  $||A||_{\infty} = \max_{i=1,...,d} \sum_{j=1}^{d} |a_{ij}|$   
The  $L_1$  - norm:  $||A||_{\infty} = \max_{j=1,...,d} \sum_{i=1}^{d} |a_{ij}|$ 

The problem becomes:

Find the minimal  $\tau > 0$  with the following property:

The matrix can be made stable/unstable by chanding elements of each row by at most  $\tau$  in the sum

	*	*	*	*	
É	*	*	*	*	j
	*	*	*	*	
	*	*	*	*	)

Note that rows of the matrix can be changed independently of each other. Every row runs over an  $L_1$  – ball independently of others.

#### We consider special sets of matrices

**Definition 1.** A family of matrices is called a *product family*, if the rows of matrices are chosen independently from given sets (*uncertainty sets*)  $F_{i}$ , i = 1, ..., d.

Example 2. A family of 3x3-matrices. The uncertainty sets are

For the first row  $a_1$ :  $F_{1} = \{(0.5, 0.2, 0.2); (0.4, 0.3, 0.2); (0.6, 0.1, 0.2); (0.55, 0.25, 0.15)\}$ For the second row  $a_2$ :  $F_2 = \{(0, 2, 1)\};$  $F_1$ For the third row  $a_3$ :  $F_3 = \{(0.4, 0.1, 2); (1, 5, 0)\}$ (0.5, 0.2, 0.2)(0.4, 0.3, 0.2)(0.6, 0.1, 0.2)(0.55, 0.25, 0.15) $\mathbf{A} = \begin{bmatrix} * & * & * \\ * & * & * \end{bmatrix}$ (0, 2, 1) $F_2$ (0.4, 0.1, 2) $F_3$ (1, 5, 0)We obtain the family M of 4x1x2 = 8 matrices

We have minimized the spectral radius over the set of eight matrices

Suppose we have four rows for each line 1, 2, and 3.

In this case we have 4x4x4 = 64 matrices. We choose one with the smallest leading eigenvalue.

# **Curse of dimensionality**

If we have d = 50 and just TWO lines in each uncertainty set, then the total number of matrices is  $2^{50} > 10^{15}$ .

Moreover, the set of rows may be polyhedral (a subset of  $\mathbb{R}^d$  defined by a system of linear inequalities).

One needs to apply some optimization technique to minimize the spectral radius over a set of matrices

#### **Product families with row uncertainties**

V.Kozyakin (2004) V.Blondel, Y.Nesterov (2009) Y.Nesterov, V.P. (2013)

#### **Applications:**

- Leontief model (mathematical economics)
- Structured population dynamics, mathematical ecology
- Spectral graph theory
- Asyncronouos systems

#### **Optimizing the spectral radius for product families**

Studied in: Y.Nesterov, V.P. (2013), V.P. (2015)

The spectral simplex method

Definition 2. A *one-line correction*, of a matrix is a replacement of one of its lines. Example 3. A correction of the first line. We replace the row  $a_1$  by some row  $a'_1$ .

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \longrightarrow A' = \begin{pmatrix} a'_{11} & a'_{12} & a'_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$
$$a'_{1} = (a'_{11}, a'_{12}, a'_{13})$$

Theorem 2. Let M be a product family of strictly positive matrices,  $F_1, ..., F_d$  be uncertainty sets. For every  $A \in M$  with the leading eigenvalue  $\lambda$  and eigenvector v, we have

a) If there is  $a'_i \in F_i$  such that  $(v, a'_i) > (v, a_i)$ , then after the one-line correction we have

$$\lambda_{\max}(A') > \lambda_{\max}(A)$$

b) If the matrix A is maximal in each row with respect to v, i.e.,

 $(v, a_i) = \max_{a'_i \in F_i} (v, a'_i), \quad i = 1, ..., d, \text{ then}$ 

$$\lambda_{\max}(A) = \max_{A' \in M} \lambda_{\max}(A')$$

## The spectral simplex method

Initialization. Take an arbitrary matrix  $A_1 \in M$ .

Main loop. We have a matrix  $A_k$  and its leading eigenvector  $v_k > 0$ .

For every i = 1, ..., d do:

Step *i*. Find  $a'_i = \underset{b_i \in F_i}{\operatorname{arg max}} (v, b_i)$ . If  $a'_i = a_i$ , then set  $A_{k+1} = A_k$  and go to the step i+1.

Otherwise, we have  $(v, a'_i) > (v, a_i)$ .

Make the one-line correction in the *i*th line. Theorem 3 implies that  $\rho(A'_k) > \rho(A_k)$ . Put  $A_{k+1} = A'_k$ . We have  $\rho(A_{k+1}) > \rho(A_k)$ . Compute the leading eigenvector  $v_{k+1} > 0$  of  $A_{k+1}$ . Go to step i = 1.

If the *d*th step is over, then END.

Theorem 3. For strictly positive matrices, the spectral simplex method is well-defined, does not cycle, and finds the solution within finite time.

Theorem 3. For strictly positive matrices, the spectral simplex method is well-defined, does not cycle, and finds the solution within finite time.

In many problems, the matrices are sparse. In this case we are in trouble.

- The leading eigenvector v of a matrix A may not be unique.
- The spectral radius is not strictly increasing with iteration, but just non-decreasing

#### The algorithm may cycle.

Example 4.  $F_1 = \{(0.5, 1, 0.5), (0.5, 0.5, 1)\}, F_2 = \{(0, 1, 0)\}, F_2 = \{(0, 0, 1)\}$ 

$$\begin{pmatrix} 0.5 & 1 & 0.5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 0.5 & 0.5 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

For sparse matrices, the algorithm cycles very often.

Theorem 4. Assume a nonnegative matrix A has a simple leading eigenvector  $v \ge 0$ , ||v||=1. Then after an arbitrary one-line correction such that  $(v, a'_i) > (v, a_i)$ , the matrix A' possesses the same property.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \longrightarrow A' = \begin{pmatrix} a'_{11} & a'_{12} & a'_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

 $a'_1 = (a'_{11}, a'_{12}, a'_{13})$  such that  $(v, a'_1) > (v, a_1)$ .

Theorem 5. If the initial matrix  $A_1$  of the spectral simplex method has a simple leading eigenvector, then all matrices in all iterations possess the same property, and the algorithm does not cycle.

How to choose  $A_1$  to possess a unique leading eigenvector ?

For instance to take the *k*th row of  $A_1$  to be the arithmetic mean of all rows from the uncertainty set  $F_k$ , for each k = 1, ..., d.

#### The numerical efficiency of the spectral simplex method

Table 1 The number of iterations for maximizing the	d/n	2	5	10	50	100
spectral radius of positive $d \times d$	5	3	6	8	10	13
matrices	10	7	12	14	18	23
	50	29	48	58	92	109
	100	56	99	131	197	213
The sets $\mathcal{F}_i$ are finite, each has $n$ elements	500	274	542	701	884	1034

For d = 100, n = 2, we have the 100-dimensional Boolean cube.

The number of vertices is  $2^{100}$ . However, the algorithm performs only 56 one-line corrections.

For d = 10, n = 100, the set M contains  $100^{10} = 10^{20}$  matrices. The algorithm performs 23 iterations.

For d = 100, n = 100, the set M contains  $10^{200}$  matrices. The algorithm performs 213 iterations.

#### The numerical efficiency of the spectral simplex method

Table 2         The number of           iterations for minimizing the	d/n	2	5	10	50	100
spectral radius of sparse $d \times d$	5	4	8	13	17	27
matrices	10	9	19	27	31	51
	50	37	56	79	131	199
Each set $\mathcal{F}_i$ has <i>n</i> elements	100	73	127	178	297	343
	500	404	742	1101	1384	1934

#### The classical simplex method (for linear programming, G.Dantzig, 1947).

LP problem: 
$$\begin{cases} (c, x) \rightarrow \max \\ (a_i, x) \leq b_i, \quad i = 1, ..., N \end{cases}$$

Step-by-step increasing of the objective function (c, x) going along the edges of the polyhedron

$$G = \{ (a_i, x) \le b_i, i = 1, ..., N \}.$$



In practice, converges extremely fast. G.Dantzig believed that the number of steps is linear in N and d.

1972. V.Klee and G.Minty constructed an example with  $2^N$  iterations.

In average, the number of iteration is indeed linear in N and d (S.Smale, 1983).

What is the theoretical complexity of the **Spectral simplex method**?

Theorem 6. For a product family M of strictly positive matrices, there are constants C > 0,  $q \in (0,1)$ , such that

$$\left|\rho(A_{N}) - \rho(A_{*})\right| \leq C q^{N},$$

where  $A_*$  is the optimal matrix,  $A_N$  is the matrix obtained in the Nth iteration of the spectral simplex method.

Corollary for the closest unstable matrix problem:

$$\begin{cases} \|X - A\|_{\infty} \to \min \\ X \ge A, \ \rho(X) = 1 \end{cases}$$
(25)

To characterize the optimal solution X we use notation  $E_k = e e_k^T$ , the matrix with the kth column of ones and all other elements are zeros.

**Theorem 5** The absolute minimum of the problem (24) is attained at the matrix

$$X = A + \tau E_k, \tag{26}$$

where  $\tau$  is reciprocal to the biggest component of the vector  $(I - A)^{-1}e$  and k is the index of that component. The value of the problem is equal to  $\tau$ .

Conclusion: to make the matrix unstable one needs to change its elements by the same number in one column!

The structure of the closest STABLE matrix is more complicated It can be found by the spectral simplex method.

The spectral simplex method works equally well for both minimization and maximization problem.

To find the closest stable matrix, one fix  $\tau > 0$ , solves the problem

 $\rho(\mathbf{X}) \to \min \\ \left\| \left\| \mathbf{X} - \mathbf{A} \right\|_{\infty} = \tau$ 

and the find the smallest  $\tau$  by bisection.

### **Applications:**

## **Optimizing the spectral radius of a graph**

Problem 1. Let us have a set of vertices  $V = \{g_1, \ldots, g_d\}$ , and for any vertex  $g_i$  a finite family of subsets  $\mathcal{V}_i$  of the set V be given. Maximize/minimize the spectral radius of a graph such that for every *i* the set of incoming edges (more precisely, the set the corresponding adjacent vertices) for the vertex  $g_i$  belongs to  $\mathcal{V}_i$ ,  $i = 1, \ldots, d$ .

Problem 2. We are given a set of vertices  $g_1, \ldots, g_d$  and a set of nonnegative integers  $n_1, \ldots, n_d$ . The problem is to find graphs with the largest and the smallest spectral radius among all directed graphs such that the number of incoming edges for any vertex  $g_i$  is equal to  $n_i$ ,  $i = 1, \ldots, d$ .

# Story 4

# Positive systems in the problems of Mathematical Economy and Population Dynamics

## **Applications:**

# The Leontief model: how to make the economy productive



Wassily Wassilievich Leontief (1906 - 1999) Василий Васильевич Леонтьев

- 1906 born in the family of W.Leontief (from an old-believer Russian orthodox family) and Genya Leontief (Becker) from a rich Jewish merchant family from Odessa.
  - 1924 Masters degree in Economics, University of Leningrad (St. Petersburg).
     Was persecuted and detained several times by Soviet authorities.
     1925 was allowed to leave Soviet Union
- 1932 1975 affiliated with Harward, from 1975 is with the New York University.
  - 1973 the Nobel Prize in Economics.

#### The Leontief input-output model (1966, Nobel Prize 1973)

expresses inter-industry relationships in linear algebraic terms.

Suppose the economy has d sectors.

The *i*th sector produces  $x_i$  units of some single homogeneous good.

To produce one unit it consumes  $a_{ij}$  units from sector *j*.

The final demand:

In addition, each sector must leave  $b_i$  units of its output to consumers

For every i = 1, ..., d, we have the equality

$$x_i = b_i + \sum_{j=1}^d a_{ij} x_j$$

Around 1949, Leontief used the primitive computer systems at Harvard to model data provided by the U.S. Bureau of Labor Statistics .

He divided the U.S. economy into d = 500 sectors.

$$x_i = b_i + \sum_{j=1}^d a_{ij} x_j, \quad i = 1, ..., d.$$

In the matrix form:  $A = (a_{ij}), \quad i, j = 1, ..., d$ :

$$(1) x = Ax + b$$

**Definition 1.** The economy is productive if it is able to provide any final demand.

#### Equation (1) has a nonnegative solution x for every nonnegative b.

Theorem 6. (W.Leontief). The economy is productive if and only if  $\rho(A) < 1$ .

#### Productivity of the economy in the Leontief model

The economy is productive  $\Leftrightarrow \lambda_{max}(A) < 1$ .

Basically, the smaller  $\lambda_{\text{max}}$  (A) the better.

$$x = \sum_{k=0}^{\infty} A^k b$$
 (reducing  $\lambda_{\max}$  makes x smaller)  $\Rightarrow$ 

The economy provides the necessary final demand b using less amount of resources x.

What to do if  $\lambda_{\max}(A) \ge 1$ ? or  $\lambda_{\max}(A)$  is very close to one ?

Example 5. We have three sectors,  $x = (x_1, x_2, x_3)$  is the value of the production.

$$\mathbf{A} = \begin{pmatrix} 0.5 & 0.2 & 0.2 \\ 0 & 0.3 & 0.7 \\ 0.5 & 0 & 0.6 \end{pmatrix}$$

 $\lambda_{\text{max}}(A) = 1 \implies$  The economy is not productive

Suppose we have a choice between four technologies producing the first good. Each technology has its own requirement of resources  $\Leftrightarrow$  has its own row  $a_1 = (a_{11}, a_{12}, a_{13})$ 

 $a_1 = (0.5, 0.2, 0.2); (0.4, 0.3, 0.2); (0.6, 0.1, 0.2); (0.55, 0.25, 0.15)$ 

$$A = \begin{pmatrix} * & * & * \\ 0 & 0.3 & 0.7 \\ 0.5 & 0 & 0.6 \end{pmatrix} (0.55, 0.25, 0.15) \implies \lambda_{\max} (A) = 1.008$$
  
$$A = \begin{pmatrix} * & * & * \\ 0 & 0.3 & 0.7 \\ 0.5 & 0 & 0.6 \end{pmatrix} (0.55, 0.25, 0.15) \implies \lambda_{\max} (A) = 0.988$$
  
$$A = \begin{pmatrix} * & * & * \\ 0 & 0.3 & 0.7 \\ 0.55, 0.25, 0.15 \end{pmatrix} \implies \lambda_{\max} (A) = 1.019$$

The third technology,  $A = \begin{pmatrix} 0.6 & 0.1 & 0.2 \\ 0 & 0.3 & 0.7 \\ 0.5 & 0 & 0.6 \end{pmatrix} \Rightarrow \lambda_{max} (A) = 0.988 \Rightarrow$  the economy is productive

#### 0. *Intro*: What are the 'Matrix Population Models'?



#### 0. *Intro*: What are the 'Matrix Population Models'?



Hans J. Math. Biol. 44, 450–462 (2002)

**Applications of Perron–Frobenius theory to population dynamics** 



**Theorem 4.4.** Let P, T and F satisfy the conditions of Theorem 4.2. For  $s > \rho(T)$  define

$$q(s) = \rho(F(I - T/s)^{-1})/s.$$
(12)

Then q(s) > 0. Let P(s) = T + F/q(s). Then its growth rate,  $\rho(P(s))$ , is s, and its net reproductive rate is

$$R_0(s) = R_0/q(s).$$

Further, one of the following holds:

 $1 = s = R_0(s), \text{ or } 1 < s < R_0(s), \text{ or } 0 < R_0(s) < s < 1.$  (13)

(1) and (2), and the net reproductive rate  $\rho(Q)$ , where  $Q = F(I - T)^{-1}$ , by  $R_0$ . Then one of the following holds:

$$r = R_0 = 1$$
, or  $1 < r \le R_0$ , or  $0 \le R_0 \le r < 1$ . (9)

*If*  $R_0 > 0$ *, then* 

$$\rho(T + F/R_0) = 1. \tag{10}$$

# Projection Matrix: general case



Problem. Let A be a nonnegative matrix. How to verify that  $\lambda_{max} > 1$  without computing  $\lambda_{max}$ ?

Indicator. A function f(A) is an indicator if  $\lambda_{max} > 1 \iff f(A) > 1$ .

Problem (D.Logofet, 2002). Does the function  $f(A) = 1 - \det(I - A)$  possess the indicator property for projection matrices ?

For some cases it was proved by Logofet and Klochkova in 2004.

This is equivalent to the property  $\lambda_2 \leq 1$ , where  $\lambda_2$  is the second largest REAL eigenvalue.

Theorem. This property always holds for matrices of the form A = T + F, where T is a substochastic matrix, F is a rank-one matrix.

Theorem. If A is a rank-one correction of T, and A, T  $\geq 0$ , then  $\lambda_2(A) \leq \lambda_{\max}(T)$ .

Corollary. For every matrix of the form A = T + F, where rank F =1, we have  $\lambda_2(A) \leq \lambda_{\max}(T)$ . If  $\lambda_{\max}(T) \leq 1$ , in particular, if T is substochastic, then  $\lambda_{\max}(A) > 1 \iff f(A) = 1 - \det (I - A) > 1$ .

Corollary. If *A* is a rank-one correction of a matrix with spectral radius  $\leq 1$ , then the function  $f(A) = \det(A-I)$ possesses the indicator property, i.e.,  $\lambda_{max}(A) > 1 \iff f(A) > 1$ .

Note that the projection matrix L = T + A is a one-line correction (since F has one nonzero line) of a substochastic matrix T, for which  $\lambda_{max}(T) \leq 1$ . Therefore,

Theorem. For every projection matrix, the function  $f(L) = \det (L-I)$ possesses the indicator property, i.e.,  $\lambda_{\max}(L) > 1 \iff 1 - \det (I-L) > 1$ .

Thank you!