Order-Stable Solutions of Linear System of Equations

Horacio Tellez Perez

University of Mons, Belgium
horacio.tellezperez@umons.ac.be

1 Introduction

Linear systems of equations are maybe the most used framework to solve dynamic systems, or at least approximate a model when exactly solving is out of the question. Not only in physics but biology, sociology and a recurrent tool on general mathematics. An example of practical use is [PW20].

Here we propose a concept related to solutions of such systems. We show also how to compute some numerical constants that allow for this new notion to be found in tractable time.

Suppose you have an square matrix where the diagonal is all zeroes: $M = \begin{bmatrix} 0 & 2 & 3 \\ 4 & 0 & 6 \\ 7 & 8 & 0 \end{bmatrix}$. If you have three non null positive real numbers a, b, c then this setting defines a linear equation system:

$$a * \mathbb{1} + b * M = c * (1, ..., 1), \tag{1}$$

which can have or not a solution. A sufficient condition is for example $\left|\frac{b}{a}\right| \|M\| < 1$, where $\|.\|$ is some norm on the space of matrices.

In any case suppose you have several solutions $X_1, ...X_n$ and you are interested on knowing if they are ordered in a similar fashion. What do we mean by order ? Let

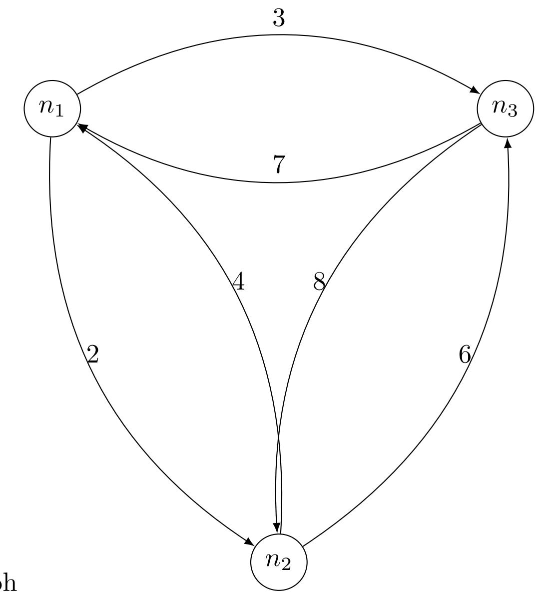
- $X_1 = (1, 2, 3)$;
- $X_2 = (3, 33, 34)$; and
- $X_3 = (33, 35, 34)$

be some solutions. Then X_1 and X_2 have composants that induce the same order on the set $\{1, 2, 3\}$ but X_3 does not. Let's formalize.

Definition 1. *Let $x, y \in \mathbb{R}^n$ be vectors, we say that x, y are rank-equivalent if the following hold:*

$$\forall 1 \leq i, j \leq n, x_i < x_j \Leftrightarrow y_i < y_j.$$

From a pure mathematical point of view this notion is not more important than one other. From a practical point of view we can have several applications however. A matrix such a one defined previously can be viewed as an interaction matrix between the nodes of a graph.



The weighted adjacency matrix of this graph is the matrix M discussed above.

Under this assumptions the a, b are the "potential" of nodes, that is a is how well we take into account the quality of importance of a node and b is the importance that we give to the interaction between different nodes. A solution to the system in 1 is the a weight given to each node so the interactions of the graph in 1 are stable.

If your aim is to find such weights, then an interest that easily come along is their order, which node has a higher weight, and if possibly, an ordering of every node. Of course, once you have a solution then this order is explicit.

2 Results

The more challenging task is to know how parameters a, b, c influence this order in the composants of the solution. First of all, the c parameter is clearly unimportant for this question. If x_c is a solution to 1 and $x_{c'}$ a solution to

$$a * \mathbb{1} + b * M = c' * (1, ..., 1), \tag{2}$$

then (recall all parameters are non zero and positive) $x_c = \frac{c}{c'} x_{c'}$ and the solutions are clearly rank-equivalent. Thus, we assume $c = 1$ for the rest of the paper. The system looks like this now:

$$a * \mathbb{1} + b * M = (1, ..., 1).$$

Second reduction that we can make, non zero parameters still, is to divide by a, we get

$$\mathbb{1} + \frac{b}{a} * M = (1, ..., 1),$$

and putting $t = \frac{b}{a}$ we find a sole parameter that can influence the ranking of vector, this new system is what we are interested in:

$$\mathbb{1} + t * M = (1, ..., 1), \tag{3}$$

We have an immediate result.

Theorem 1 (Convergence). *Let M be an $N \times N$ matrix whose diagonal elements are zero. Let $b, c \in \mathbb{R}$. Consider the following equation with real-number parameter a:*

$$(a * \mathbb{1} - b * M) \cdot X = \begin{bmatrix} c \cdots c \end{bmatrix}^\intercal . \tag{4}$$

Then there exists $a^* \in \mathbb{R}$ such that

- *for every $a \geq a^*$, the equation (4) has a unique solution; and*
- *for all $a_1, a_2 \geq a^*$, if X_1 and X_2 are solutions to (4) for parameters a_1 and a_2 respectively, then X_1 and X_2 are rank-equivalent.*

Moreover, such a bound a^* can be computed in polynomial time in the size of N.

Proof. Clearly, if x_c is the solution to

$$(a * \mathbb{1} - b * M) \cdot X = c * \begin{bmatrix} 1 \; 1 \cdots \; 1 \end{bmatrix}^\intercal$$

with $c > 0$, and x_1 is the solution to

$$(a * \mathbb{1} - b * M) \cdot X = \begin{bmatrix} 1 \; 1 \cdots \; 1 \end{bmatrix}^\intercal ,$$

then $x_c = c * x_1$. Since x_c and x_1 are rank-equivalent, we can fix $c = 1$ without loss of generality. Define a' as follows:

$$a' := 1 + |b| * (N - 1) \left(\max_{1 \leq i,j \leq N} |M_{ij}| \right) .$$

By the Levy-Desplanques theorem, for every $a \geq a'$, the matrix $a * \mathbb{1} - b * M$ is invertible. For every $a \geq a'$, we write x^a for the unique solution of the equation

$$(a * \mathbb{1} - b * M) \cdot X = \begin{bmatrix} 1 \; 1 \cdots \; 1 \end{bmatrix}^\intercal .$$

For $a \geq a'$, define $\delta_{ij}(a) := x_i^a - x_j^a$, where x_i^a is the ith coordinate of x^a.

We will show that there is a value $a^* \geq a'$ such that for all $a_1, a_2 \geq a^*$, for all $1 \leq i < j \leq N$, $\delta_{ij}(a_1)$ and $\delta_{ij}(a_2)$ have the same sign, which implies that x^{a_1} and x^{a_2} are rank-equivalent.

Define the following matrix $S_{|i}$ in function of a, for $1 \leq i \leq N$:

$$(S_{|i}(a))_{\ell k} = \begin{cases} (a * \mathbb{1} - b * M)_{\ell k} & \text{if } k \neq i \\ 1 & \text{if } k = i \end{cases} \tag{5}$$

That is, $S_{|i}$ *is obtained from* $a * \mathbb{1} - b * M$ *by replacing the* i*th column with* $\left[1\ 1 \cdots 1\right]^\intercal$*. By Cramer's Rule, we have*

$$x_i^a = \frac{\det\left(S_{|i}(a)\right)}{\det\left(a * \mathbb{1} - b * M\right)}, \tag{6}$$

and consequently

$$\delta_{ij}(a) = \frac{\det\left(S_{|i}(a)\right) - \det\left(S_{|j}(a)\right)}{\det\left(a * \mathbb{1} - b * M\right)}. \tag{7}$$

Since the determinants $\det(\cdot)$ *are polynomial expressions, the following are all polynomials of degree at most* N*:*

- $p_i(a) = \det\left(S_{|i}(a)\right)$*;*
- $p_j(a) = \det\left(S_{|j}(a)\right)$*; and*
- $p(a) = \det\left(a * \mathbb{1} - b * M\right)$*.*

If $a \geq a'$*, then since the polynomial* $p(a)$ *does not vanish, the sign of* $\delta_{ij}(a)$ *is fully determined by the expression*

$$\det\left(S_{|i}(a)\right) - \det\left(S_{|j}(a)\right). \tag{8}$$

Let $p_{ij} = p_i - p_j$ *for* $1 \leq i < j \leq N$*. We determine* a^* *such that for all* $a \geq a^*$ *and for all* $1 \leq i < j \leq N$*, the sign of* $p_{ij}(a)$ *does not change (i.e., is either always positive or always negative). Note that the sign of* p_{ij} *not changing is equivalent to the sign of* p_{ji} *not changing, so we can assume* $i < j$ *without loss of generality.*

In the remainder of the proof, we show that the desired a^* *exists and can be computed in polynomial time in* N*. Pick* $N+1$ *real numbers* $V = \{a_1, \ldots, a_{N+1}\}$*, all greater than* a'*. Compute*

$$p_{ij}(a_k) = \det\left(S_{|i}(a_k)\right) - \det\left(S_{|j}(a_k)\right) \tag{9}$$

for all $a_k \in V$ *and* $1 \leq i < j \leq N$*. The set*

$$\{p_{ij}(a_k) \mid 1 \leq i < j \leq N, 1 \leq k \leq N+1\}$$

can be computed in polynomial time in N*, because it involves* $N(N+1)$ *determinants (i.e.,* $\det\left(S_{|i}(a_k)\right)$ *for* $1 \leq i \leq N$ *and* $1 \leq k \leq N+1$*), each of which can be computed in polynomial time in* N*. For all* $1 \leq i < j \leq N$*, the polynomial* p_{ij} *can be computed from* $\{(a_1, p_{ij}(a_1)), \ldots, (a_{N+1}, p_{ij}(a_{N+1}))\}$ *in polynomial time using, for example, Lagrange interpolation. The number of polynomials to compute is* $\frac{N(N-1)}{2}$ *(i.e., polynomially many), and each of them has at most* N *coefficients. We will represent the polynomial* p_{ij} *by its coefficients, i.e., by* $\langle (p_{ij})_0, \ldots, (p_{ij})_{n_{ij}} \rangle$ *where each* $(p_{ij})_\ell$ *is the coefficient of degree* ℓ*, and* $n_{ij} \leq N$ *is the polynomial's degree. We now define* a^*_{ij} *as*

$$a^*_{ij} := \max\left(a', 2 + \max_{0 \leq \ell \leq n_{ij}-1} \frac{-(p_{ij})_\ell}{\left|(p_{ij})_{n_{ij}}\right|}\right).$$

By Cauchy's bound on positive real roots of polynomials, if x_0 *is a root of* p_{ij}*, then* $x_0 < a^*_{ij}$*. This implies that if* $a_1, a_2 > a^*_{ij}$*, then* $p_{ij}(a_1)$ *and* $p_{ij}(a_2)$ *have the same sign (i.e., either both positive or both negative), and therefore, by definition of* δ_{ij}*, we have* $x_i^{a_1} < x_j^{a_1}$ *if and only if* $x_i^{a_2} < x_j^{a_2}$*. Finally, let*

$$a^* = \max_{1 \leq i < j \leq N} a^*_{ij},$$

which can be computed in polynomial time. By our construction, if follows that for all $a_1, a_2 > a^*$*, the solutions* x^{a_1}, x^{a_2} *exist and are rank-equivalent.*

Finally, we incidentally note that a slightly better bound is obtained by letting

$$a^* = \max\left(a', 2 + \max_{1 \leq i < j \leq N, 0 \leq \ell \leq n_{ij}-1} \frac{-(p_{ij})_\ell}{|(p_{ij})_{n_{ij}}|}\right). \tag{10}$$

This concludes the proof. □

3 Implementation

If people are patient we will explain how to accurately and quickly compute the coefficients needed in 1. Recall that coefficient of an interpolation polynomial are given by a Vandermonde matrix, that are famously ill-conditioned [Pan16].

References

[Pan16] Victor Y. Pan. How bad are vandermonde matrices? *SIAM J. Matrix Anal. Appl.*, 37(2):676–694, 2016.

[PW20] Horacio Tellez Perez and Jef Wijsen. Logic-based ranking of assertions in inconsistent aboxes. In Stefan Borgwardt and Thomas Meyer, editors, *Proceedings of the 33rd International Workshop on Description Logics (DL 2020) co-located with the 17th International Conference on Principles of Knowledge Representation and Reasoning (KR 2020), Online Event [Rhodes, Greece], September 12th to 14th, 2020*, volume 2663 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2020.