A Perron–Frobenius Theorem for Jordan Blocks for Complexity Proving

Jose Divasón[1], Sebastiaan Joosten[2], René Thiemann[3], and Akihisa Yamada[4]

1 **Universidad de La Rioja, Spain**
2 **University of Twente, the Netherlands**
3 **University of Innsbruck, Austria**
4 **National Institute of Informatics, Japan**

1 Introduction

Matrix interpretations [4] are widely used in automated complexity analysis of term rewrite systems. We consider the approach where from a matrix interpretation one extracts a nonnegative square matrix $A \in \mathbb{R}_{\geq 0}^{n \times n}$ such that $\mathcal{O}(\max_{1 \leq i,j \leq n}(A^k)_{ij} \cdot k)$ is a bound for the interpretation of a (specific) term of size k [8]. Hence, the growth rate of the entries of A^k can be used to determine an upper bound on the complexity of the rewrite system.

In this paper we concentrate on algebraic means to investigate the growth rate of A^k. We use the following notions: χ_A denotes the characteristic polynomial of matrix A, eigenvalues are the roots of χ_A, maximal eigenvalues are those of the maximum norm, and the maximum norm is the spectral radius ρ_A. A Jordan block for some eigenvalue λ is a square matrix B with $B_{ii} = \lambda$, $B_{i(i+1)} = 1$, and $B_{ij} = 0$ otherwise. We write $B(s, \lambda)$ to indicate that a Jordan block has size s and eigenvalue λ. A diagonal composition J of Jordan blocks is a Jordan normal form of A if $A = PJP^{-1}$ for some invertible matrix P; J is unique up to permutation of the Jordan blocks, so we call them *the Jordan blocks of A*.

Jordan blocks are an important utility to characterize matrix growth, cf. Theorem 1. This theorem was formalized in Isabelle/HOL [10, 12] when formalizing the complexity criteria provided by Neurauter et al. [7, 9]. It was used in CeTA [13] until year 2017 in order to check matrix growth rates.

▶ **Theorem 1** (Matrix Growth via Algebraic Methods). *Let $A \in \mathbb{C}^{n \times n}$.*

- *The entries in A^k are polynomially bounded in k iff $\rho_A \leq 1$.*
- *The entries in A^k are bounded by $\mathcal{O}(k^d)$ iff $\rho_A \leq 1$ and for all eigenvalues λ of A with $|\lambda| = 1$ and for all Jordan blocks $B(s, \lambda)$ of A, $s \leq d + 1$.*

Using Theorem 1 it seems easy to decide the asymptotic growth rate of a matrix: one can compute all eigenvalues, compute their maximal norms to test $\rho_A \leq 1$, and then compute all Jordan blocks of A for eigenvalues with norm 1 and determine their maximal size.

The problem is that in general the eigenvalues of a matrix are quite complex to compute. Even if all matrix entries are integers or rational numbers, the eigenvalues are complex algebraic numbers. Hence, applying Theorem 1 directly requires expensive algebraic number arithmetic.

The aim of this work is to reduce the complexity of applying Theorem 1 for an important class of matrices: non-negative real square matrices $\mathbb{R}_{\geq 0}^{n \times n}$.

Several properties of non-negative real matrices are studied by Perron and Frobenius [5, 11]. We already formalized several of them in developing an efficient certifier for checking matrix growth, for instance Theorem 2 [2].

Proceedings of the 16th International Workshop on Termination;
Editor: Salvador Lucas

▶ **Theorem 2.** *The characteristic polynomial χ_A of $A \in \mathbb{R}^{n\times n}_{\geq 0}$ can be factored into*

$$\chi_A(x) = f(x) \cdot \prod_{k\in K} (x^k - \rho_A^k)$$

where K is a non-empty multiset of positive integers, and f is a polynomial whose complex roots have a norm strictly less than ρ_A.

Note that Theorem 2 has at least three implications. First, ρ_A itself is an eigenvalue, so $\rho_A \leq 1$ is the same as demanding that χ_A has no *real* eigenvalue above 1—a criterion which can be efficiently decided by Sturm's method. Second, it permits us to replace the computation of *all* eigenvalues in Theorem 1 by a computation which only involves roots of unity of degree at most n. Third, ρ_A has the maximum algebraic multiplicity among all maximal eigenvalues.

For a precise and efficient growth rate analysis, in this paper we further prove a conjecture in [2, Section 7]: ρ_A also has the largest Jordan blocks among all maximal eigenvalues. Previously we only had a partial result for matrices of dimension up to 4.

We also derive some consequences in Section 3.

2 Largest Jordan Blocks for Non-Negative Real Matrices

The following Theorem 3 is the key new result of this paper. It has been formalized in Isabelle/HOL and will be available in the archive of formal proofs [3, for Isabelle 2018].

▶ **Theorem 3.** *Let $A \in \mathbb{R}^{n\times n}_{\geq 0}$ and λ be a maximal eigenvalue of A. For every Jordan block $B(s, \lambda)$ of A, there exists a Jordan block $B(t, \rho_A)$ of A with $t \geq s$.*

Proof. W.l.o.g. we assume $\rho_A = 1$, since otherwise we can just perform a scalar multiplication of A by $\frac{1}{\rho_A}$ (The case $\rho_A = 0$ is trivial).

Let $A = PJP^{-1}$ be the Jordan decomposition of A. Let e_i denote the i-th unit-vector. Let E be the set of maximal eigenvalues of A, i.e., eigenvalues λ with $|\lambda| = 1$. Let m be the size of a largest Jordan block among all blocks with eigenvalues in E. By Theorem 2 all eigenvalues λ with $|\lambda| = 1$ are roots of unity, i.e., we have a positive integer d_λ such that $\lambda^{d_\lambda} = 1$.

Define

$$c = \prod_{0\leq i<m-1} (m-1-i) \qquad D = \prod_{\lambda\in E} d_\lambda \qquad K_\ell\ k = D \cdot k + \ell + m - 1$$

It is easy to prove $c > 0$. Moreover, if $\lambda \in E$, then $\lambda^{K_\ell\ k}$ does not depend on k.

$$\lambda^{K_\ell\ k} = (\lambda^D)^k \cdot \lambda^\ell \cdot \lambda^{m-1} = 1^k \cdot \lambda^\ell \cdot \lambda^{m-1} = \lambda^\ell \cdot \lambda^{m-1}.$$

Moreover for any Jordan Block B of size s with eigenvalue λ of A and any valid position of B we get the following limit:

$$\frac{c(B^{K_\ell\ k})_{ij}}{(K_\ell\ k)^{m-1}} = \frac{c\binom{K_\ell\ k}{j-i}\lambda^{K_\ell\ k+i-j}}{(K_\ell\ k)^{m-1}} \xrightarrow{k\to\infty} \begin{cases} \lambda^\ell & \text{if } m = s = j,\ i = 1, \text{ and } \lambda \in E \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

The equality in (1) is just the closed form for powers of Jordan blocks, and we shortly illustrate by a case analysis why the limit is correct.

- If $\lambda \notin E$ then there is an exponential decrease in the $\lambda^{\cdots k \cdots}$ expression whereas the remaining expression grows at most polynomial. So the limit will be 0.
- If $\lambda \in E$, but $m = j \wedge i = 1$ does not hold, then the norm of the numerator will be a polynomial of smaller degree than the polynomial in the denominator. So, again the limit will be 0.
- Finally, if $m = j$, $i = 1$, and $\lambda \in E$, then both the numerator and the denominator will be polynomials of the same degree, and c has been defined in a way that the limit will be precisely λ^ℓ.

The next step is to lift (1) from single Jordan blocks to the complete Jordan normal form J. Let us define I as the set of last-row indices of those Jordan blocks of J which have size m and a maximal eigenvalue. So, in particular I is non-empty. Then from (1) we immediately get the following consequence where $\lambda_i = J_{ii}$ is the eigenvalue corresponding to row i.

$$\frac{c(J^{K_\ell\ k})_{ij}}{(K_\ell\ k)^{m-1}} \overset{k\to\infty}{\longrightarrow} \begin{cases} \lambda_j^\ell & \text{if } j \in I \wedge i = j-(m-1) \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Define

$$\begin{aligned} i &= \text{some element of } I \\ j &= \text{some index } j \text{ such that } P_{j(i-(m-1))} \neq 0 \\ v &= |Pe_i| \end{aligned}$$

Then

$$\begin{aligned} \frac{c(A^{K_\ell\ k}v)_j}{(K_\ell\ k)^{m-1}} &= \frac{c(|A^{K_\ell\ k}| \cdot |v|)_j}{(K_\ell\ k)^{m-1}} = \frac{c(|A^{K_\ell\ k}| \cdot |Pe_i|)_j}{(K_\ell\ k)^{m-1}} \\ &\geq \frac{c|A^{K_\ell\ k}Pe_i|_j}{(K_\ell\ k)^{m-1}} = \frac{c|PJ^{K_\ell\ k}e_i|_j}{(K_\ell\ k)^{m-1}} \\ &\overset{k\to\infty}{\longrightarrow} |P_{j(i-(m-1))} \cdot \lambda_i^\ell| = |P_{j(i-(m-1))}| > 0 \end{aligned}$$

Thus, there exists some $b > 0$ and such that for all ℓ and all sufficiently large k:

$$\frac{c(A^{K_\ell\ k}v)_j}{(K_\ell\ k)^{m-1}} \geq b \tag{3}$$

Define

$$\begin{aligned} u &= P^{-1}v \\ a_i &= P_{j,i-(m-1)} \cdot u_i \qquad \text{for } j \text{ as above and arbitrary } i \end{aligned}$$

Then by using (2) conclude

$$\begin{aligned} \frac{c(A^{K_\ell\ k}v)_j}{(K_\ell\ k)^{m-1}} &= Re\left(\frac{c(A^{K_\ell\ k}v)_j}{(K_\ell\ k)^{m-1}}\right) = Re\left(\frac{c(A^{K_\ell\ k}Pu)_j}{(K_\ell\ k)^{m-1}}\right) \\ &= Re\left(\frac{c(PJ^{K_\ell\ k}u)_j}{(K_\ell\ k)^{m-1}}\right) \overset{k\to\infty}{\longrightarrow} Re\left(\sum_{i\in I} a_i\lambda_i^\ell\right) \end{aligned} \tag{4}$$

By combining (4) with (3) we arrive at

$$Re\left(\sum_{i\in I} a_i\lambda_i^\ell\right) > 0 \tag{5}$$

The theorem is equivalent to the statement $\lambda_i = 1$ for some $i \in I$, so let us assume to the contrary that $\lambda_i \neq 1$ for all $i \in I$. From (5) we conclude

$$0 < Re\left(\sum_{\ell=0}^{D-1} \sum_{i \in I} a_i \lambda_i^\ell\right)$$

and hence

$$\begin{aligned} 0 \neq \sum_{\ell=0}^{D-1} \sum_{i \in I} a_i \lambda_i^\ell &= \sum_{i \in I} a_i \sum_{\ell=0}^{D-1} \lambda_i^\ell = \sum_{i \in I} a_i \frac{D}{d_{\lambda_i}} \cdot \sum_{\ell=0}^{d_{\lambda_i}-1} \lambda_i^\ell \\ &= \sum_{i \in I} a_i \frac{D}{d_{\lambda_i}} \cdot \frac{\lambda_i^{d_{\lambda_i}} - 1}{\lambda_i - 1} = \sum_{i \in I} a_i \frac{D}{d_{\lambda_i}} \cdot 0 = 0 \end{aligned}$$

where the formula for the geometric sum can only be applied since $\lambda_i \neq 1$. By this contradiction we have finished the proof. ◀

3 Consequences of Theorem 3

We can immediately connect Theorem 3 with Theorem 1 in order to get an improved complexity criterion in the form of Algorithm 1. Here, I is the identity matrix.

The big advantage of this algorithm is that it only employs arithmetic operations, i.e., in particular if A is a matrix with rational entries, then it only involves rational arithmetic.

Algorithm 1: Efficient Certification of $\max_{1 \leq i,j \leq n} (A^k)_{ij} \in \mathcal{O}(k^d)$.

Input: $A \in \mathbb{R}_{\geq 0}^{n \times n}$ and degree d.
Output: Accept or assertion failure.
Assert $\{x \in \mathbb{R}. \chi_A(x) = 0, x > 1\} = \emptyset$ via Sturm's method
Compute m as the multiplicity of root 1 of χ_A
if $m \leq d+1$ **then** Accept
Assert that all Jordan blocks $B(s, 1)$ of A have a size $s \leq d+1$. This can be decided by checking that the kernel dimension of $(A - I)^{d+1}$ is m
Accept

To measure improvements in practice, we extracted all matrix interpretations from complexity proofs of the international termination and complexity competition [6] in the last three years, which amounts to the validation of the growth rate of 6,690 matrices, whose largest dimension was only 5. This low dimension keeps the overhead of algebraic number computations at a reasonable level. Still, processing all 6,690 matrices became five times faster when using Algorithm 1 instead of Theorem 1 and algebraic number computations.

Since the algorithm runs in polynomial time, it also seems to be possible to use our algorithm for synthesizing matrix interpretations; one can write a polynomial-sized SAT or SMT encoding of whether a symbolic matrix has an *a priori* fixed growth rate by a symbolic execution of Algorithm 1. The required algorithms are

- computation of characteristic polynomial χ_A
- polynomial division and polynomial GCD for applying Sturm's method
- Gauss-Jordan elimination for computing the kernel dimension, and matrix multiplication

Developers of the complexity tool TCT [1] are currently investigating this possibility.

Although our results facilitate precise estimation of the asymptotic growth rate of matrix powers, the reduction of complexity analysis to matrix powers remains approximative. Therefore, it is an important future work to generalize our formalization to *joint* spectral radius, in order to facilitate precise complexity analysis [7].

Acknowledgments

We thank Hans Zwart for discussions on Theorem 3. He came up with an alternative proof at a stage where already a large part of our proof has been formalized.

This research was supported by the Austrian Science Fund (FWF) project Y757. Jose Divasón is partially funded by the Spanish projects MTM2014-54151-P and MTM2017-88804-P. Akihisa Yamada is partially supported by ERATO HASUO Metamathematics for Systems Design Project (No. JPMJER1603), JST. The authors are listed in alphabetical order regardless of individual contributions or seniority.

References

1 Martin Avanzini, Georg Moser, and Michael Schaper. TcT: Tyrolean Complexity Tool. In *TACAS 2016*, volume 9636 of *LNCS*, pages 407–423, 2016.

2 Jose Divasón, Sebastiaan Joosten, Ondřej Kunčar, René Thiemann, and Akihisa Yamada. Efficient certification of complexity proofs: Formalizing the Perron–Frobenius theorem (invited talk paper). In *Proceedings of the 7th ACM SIGPLAN International Conference on Certified Programs and Proofs*, CPP 2018, pages 2–13. ACM, 2018.

3 Jose Divasón, Ondřej Kunčar, René Thiemann, and Akihisa Yamada. Perron–Frobenius theorem for spectral radius analysis. *Archive of Formal Proofs*, May 2016. `http://isa-afp.org/entries/Perron_Frobenius.shtml`, Formal proof development.

4 Jörg Endrullis, Johannes Waldmann, and Hans Zantema. Matrix interpretations for proving termination of term rewriting. *Journal of Automated Reasoning*, 40(2-3):195–220, 2008.

5 Ferdinand Georg Frobenius. Über Matrizen aus nicht negativen Elementen. In *Sitzungsberichte Preuß. Akad. Wiss.*, pages 456–477, 1912.

6 Jürgen Giesl, Frédéric Mesnard, Albert Rubio, René Thiemann, and Johannes Waldmann. Termination competition (termCOMP 2015). In *CADE-25*, volume 9195 of *LNCS*, pages 105–108, 2015.

7 Aart Middeldorp, Georg Moser, Friedrich Neurauter, Johannes Waldmann, and Harald Zankl. Joint spectral radius theory for automated complexity analysis of rewrite systems. In *CAI 2011*, volume 6742 of *LNCS*, pages 1–20, 2011.

8 Georg Moser, Andreas Schnabl, and Johannes Waldmann. Complexity analysis of term rewriting based on matrix and context dependent interpretations. In *FSTTCS 2008*, volume 2 of *LIPIcs*, pages 304–315, 2008.

9 Friedrich Neurauter, Harald Zankl, and Aart Middeldorp. Revisiting matrix interpretations for polynomial derivational complexity of term rewriting. In *LPAR-17*, volume 6397 of *LNCS*, pages 550–564, 2010.

10 Tobias Nipkow, Lawrence C. Paulson, and Makarius Wenzel. *Isabelle/HOL – A Proof Assistant for Higher-Order Logic*, volume 2283 of *LNCS*. Springer, 2002.

11 Oskar Perron. Zur Theorie der Matrices. *Math. Ann.*, 64:248–263, 1907.

12 René Thiemann and Akihisa Yamada. Formalizing Jordan normal forms in Isabelle/HOL. In *CPP 2016*, pages 88–99. ACM, 2016.

13 René Thiemann and Christian Sternagel. Certification of termination proofs using CeTA. In *TPHOLs'09*, volume 5674 of *LNCS*, pages 452–468, 2009.