51st IEEE Conference on Decision and Control
December 10-13, 2012. Maui, Hawaii, USA

Joint Spectral Radius of Rank One Matrices and the Maximum Cycle Mean Problem

Amir Ali Ahmadi and Pablo A. Parrilo

***Abstract*— We show that the problem of exact computation of the joint spectral radius of a finite set of rank one matrices can be reformulated as the problem of computing the maximum cycle mean in a directed graph and hence be solved efficiently.**

I. Introduction

Given a finite set of square matrices $\mathcal{A} := \{A_1, \ldots, A_m\}$, their *joint spectral radius* $\rho(\mathcal{A})$ is defined as

$$\rho(\mathcal{A}) = \lim_{k\to\infty} \max_{\sigma\in\{1,\ldots,m\}^k} \|A_{\sigma_k}\cdots A_{\sigma_2}A_{\sigma_1}\|^{1/k}, \quad (1)$$

where the quantity $\rho(\mathcal{A})$ is independent of the norm used in (1). The joint spectral radius (JSR) is a natural generalization of the spectral radius of a single matrix and it characterizes the maximal growth rate that can be obtained by taking products of arbitrary length, of all possible permutations of $A_1, \ldots, A_m$. This concept was introduced by Rota and Strang [1] in the early 60s and has since emerged in many areas of application such as stability of switched linear systems, computation of the capacity of codes, continuity of wavelet functions, convergence of consensus algorithms, and many others; see [2] and references therein. In particular, the switched linear dynamical system $x_{k+1} = A_i x_k$, $i = 1, \ldots, m$, is asymptotically stable under arbitrary switching if and only if $\rho(\mathcal{A}) < 1$.

There are several undecidability and NP-hardness results demonstrating the difficulty of computing the JSR in general [3], [4]. Nevertheless, the significance of the concept has encouraged researchers to come up with a variety of algorithms for providing lower and upper bounds on the JSR or identifying special cases where it can be computed exactly and efficiently; see e.g. [5], [6], [7], [8], [2].

The focus of this short note is on the computation of the JSR of a set of *rank one* matrices, a problem that was recently motivated and studied by Liu and Xiao in [9]. In there, the authors also study the JSR of more general sets of matrices by approximating them with a set of rank one matrices. It is claimed in [9] that if the set $\mathcal{A}$ consists only of rank one matrices, then $\rho(\mathcal{A})$ is given by the square root of the spectral radius of a matrix product of length *two* from $\mathcal{A}$. We will show that this is not the case (Example 2.1), though it is true that the JSR is always achieved by the spectral radius of a product of length at most m (Corollary 2.1). Our main contribution, however, is to show that one can efficiently compute the JSR of a set of rank one matrices exactly by using algorithms for the *maximum cycle mean problem* in graph theory [10], [11]. This is done in Section II. Some illustrating examples are given in Section III.

Amir Ali Ahmadi is with the Krannert School of Management at Purdue University and currently a Goldstine Fellow at the Department of Business Analytics and Mathematical Sciences of the IBM Watson Research Center; Email: a_a_a@mit.edu. Pablo A. Parrilo is with the Laboratory for Information and Decision Systems, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology; Email: parrilo@mit.edu.

II. Main Result

We start by recalling some basic facts about rank one matrices. A real $n\times n$ matrix A is rank one if and only if there exist real vectors x and y in $\mathbb{R}^n$ such that $A = xy^T$. The spectral radius $\rho(A)$ of such a matrix is equal to the absolute value of its only nonzero eigenvalue which is $y^T x$. The product of two rank one matrices $A_i = x_i y_i^T$ and $A_j = x_j y_j^T$ is $A_iA_j = (y_i^T x_j)x_i y_j^T$ and therefore has rank at most one. By induction, for any k, products of length k out of a set of rank one matrices have rank at most one.

Construction of $G_{\mathcal{A}}$. Given a set of rank one matrices $\mathcal{A} = \{A_1, \ldots, A_m\}$, $A_i = x_i y_i^T$, we define the graph $G_{\mathcal{A}}$ to be a complete graph with m nodes, one per matrix, and m^2 directed edges e_{ij}, $\forall i, j \in \{1, \ldots, m\}^2$, where to edge e going from node i to node j we assign a weight $w(e) = |y_i^T x_j|$.

By a cycle, we will always mean a directed cycle in what follows. A *simple* cycle is a cycle that does not visit any node more than once. We denote the set of all simple cycles in the graph $G_{\mathcal{A}}$ by C, and clearly $|C|$ is finite. Let the sequence of edges $(e_1, e_2, \ldots, e_k)$ form a cycle c. We denote the product of the weights on the edges by ρ_c, i.e., $\rho_c = \prod_{i=1}^k w(e_i)$. The *gain* $g(c)$ of the cycle is defined to be $\rho_c^{1/k}$. Let $\rho^* = \max_c g(c)$, where c ranges over all simple cycles in C; ρ^* is called the *maximum cycle gain* of $G_{\mathcal{A}}$ and a cycle c_{max} that achieves it is called a *gain maximizing cycle*.

Given the facts that we recalled about products of rank one matrices, the proof of the following lemma should be obvious and is hence omitted.

Lemma 2.1: Let $A_{\sigma_k}\cdots A_{\sigma_1}$ be a product from the set $\mathcal{A} = \{A_1, \ldots, A_m\}$. Let c be a cycle in $G_{\mathcal{A}}$ consisting of k edges that go from node A_{σ_1} to node A_{σ_k} and then back to A_{σ_1} (through a single returning edge). Then,

$$\rho(A_{\sigma_k}\cdots A_{\sigma_1}) = \rho_c.$$

Theorem 2.2: Consider a set of rank one matrices $\{A_1, \ldots, A_m\}$, and the associated complete directed graph $G_{\mathcal{A}}$ as above. Let c_{max} be a gain maximizing cycle for $G_{\mathcal{A}}$ of let l_{max} and $\rho_{c_{max}}$ respectively denote its length, and the product of the weights on its edges. The joint spectral radius is then given by the maximum cycle gain of $G_{\mathcal{A}}$, i.e.,

$$\rho(\mathcal{A}) = \rho_{c_{max}}^{1/l_{max}}.$$

Moreover, it is achieved with a finite matrix product, corresponding to the product of the matrices (nodes) in c_{max}.

Proof: Consider an arbitrary product $A_{\sigma_k} \cdots A_{\sigma_1}$ from the set $\mathcal{A}$. For some integer s, we can decompose the cycle associated to this product into the concatenation of s simple cycles $c_1, \ldots, c_s$ from the set C. We denote the multiplicity of the simple cycle c_i in this decomposition by m_i and its length by l_i. Using our notation above, we have

$$\begin{array}{rcl} \rho(A_{\sigma_1} \cdots A_{\sigma_k})^{1/k} & = & (\prod_{i=1}^{s} \rho_{c_i}^{m_i})^{1/k} \\ & = & \prod_{i=1}^{s} (\rho_{c_i}^{1/l_i})^{m_i l_i / k} \\ & \leq & \rho_{c_{max}}^{1/l_{max}}, \end{array} \qquad (2)$$

where the first equality follows from Lemma 2.1, the second equality is obvious, and the final inequality follows from the definition of $\rho_{c_{max}}$ and the fact that $\sum_{i=1}^{s} m_i l_i = k$.

Let $\mathcal{A}^k$ denote the set of all matrix products of length k. The following characterization of the JSR is well-known:

$$\rho(\mathcal{A}) \;=\; \limsup_{k\to\infty} \max_{A\in\mathcal{A}^k} \rho^{\frac{1}{k}}(A).$$

See e.g. [2, Chap. 1]. This characterization, together with (2), immediately imply that $\rho(\mathcal{A}) = \rho_{c_{max}}^{1/l_{max}}$. The proof of the latter claim should be obvious. ■

Corollary 2.1: The JSR of a set of m rank one matrices is achieved by the spectral radius of a matrix product of length at most m. (This in particular implies that the *"finiteness property*[1]*"* always holds for rank one matrices). Moreover, there will always be a JSR achieving matrix product in which no matrix appears more than once.

Proof: This simply follows from the fact that a simple cycle does not visit a node twice. ■

It is claimed in [9] that it is enough to consider the spectral radius of products of length two for computing the JSR of rank one matrices. We next give a counterexample to this claim, demonstrating that matrix products of length m in Corollary 2.1 may indeed be required.

Example 2.1: Let e_i be the ith vector of the standard basis of $\mathbb{R}^m$. Let $\mathcal{A} = \{A_1, \ldots, A_m\}$, where

$$A_1 = e_1 e_2^T, \quad A_2 = e_2 e_3^T, \quad \ldots \quad A_m = e_m e_1^T.$$

Since $\{e_1, \ldots, e_m\}$ satisfy the orthonormality condition $e_i^T e_j = \delta_{ij}$, it follows that $\rho(A_i A_j) = 0$ for all $i, j \in \{1, \ldots, m\}$, and that the only nonzero infinite products are cyclic repetitions of $A_1 A_2 \cdots A_m$. Thus, for this example the joint spectral radius is $\rho(\mathcal{A}) = 1$, and it is only achieved by products of length that are an integer multiple of m.

[1]See [2] for the definition of the finiteness property and a discussion of the well-known *finiteness conjecture*.

A. The maximum cycle mean problem

So far we have established that the computation of the JSR for a set of m rank one matrices $\mathcal{A}$ can be reduced to the task of finding the maximum cycle gain of $G_{\mathcal{A}}$. The naive algorithm for doing this would enumerate all possible cycles of length at most m and have exponential running time in m. We now show the immediate connection of this problem to the *maximum cycle mean problem* (MCMP) in graph theory which enables us to get a very efficient polynomial time algorithm for computing the JSR.

In the maximum cycle mean problem, one is given a directed graph $G(V,E)$ together with a cost function $f : E \to \mathbb{R}$ on its edges. Given a path $\sigma = (e_1, e_2, \ldots, e_k)$ of length k in the graph, its *mean weight* $m(\sigma)$ is defined as $m(\sigma) = \sum_{i=1}^{k} \frac{f(e_i)}{k}$. The *maximum cycle mean* λ^* of the graph is defined to be $\lambda^* = \max_c m(c)$, where c ranges over all cycles in the graph. (It is easy to see that the maximum can always be achieved at a simple cycle.)

Let us define a graph $\tilde{G}_{\mathcal{A}}$ that is identical to $G_{\mathcal{A}}$ in structure except that the weight on any of its edges e_{ij} is equal to the logarithm of the edge weight of e_{ij} in $G_{\mathcal{A}}$, i.e., is equal to $\log|y_i^T x_j|$. Then, in view of Theorem 2.2, it is not difficult to see that the JSR is given by

$$\rho(\mathcal{A}) = e^{\lambda^*/k^*},$$

where λ^* is the maximum cycle mean of $\tilde{G}_{\mathcal{A}}$ and k^* is the length of a maximizing cycle.

Remark 2.1: In [10], Karp gave an efficient dynamic programming based algorithm for computing λ^* and extracting a maximizing cycle[2]. The running time of Karp's algorithm is $O(|N||E|)$. From this it follows that the JSR of a set of m rank one $n \times n$ matrices can be computed in $O(m^3 + m^2 n)$. More efficient algorithms for MCMP have also appeared since [12].

Remark 2.2: Since the algorithms for MCMP work for arbitrary directed graphs and not just for complete ones, our approach can be immediately extended to the problem of analyzing stability of rank one linear systems under *constrained switching*.

III. Examples

We present now a few simple examples:

Example 3.1: This example was considered in [9, Ex. 1]. Let

$$A_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}^T, \qquad A_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}^T.$$

For this factorization, the corresponding matrix of weights (i.e., the adjacency matrix of $\tilde{G}_{\mathcal{A}}$), is

$$\log \begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -\infty & \log 2 \\ 0 & 0 \end{bmatrix}.$$

The maximum mean cycle is clearly given by $\{1, 2\}$, with cycle mean $\frac{\log 2}{2}$, and thus the joint spectral radius is $\rho(\{A_1, A_2\}) = \exp(\frac{\log 2}{2}) = \sqrt{2}$.

[2]Karp's algorithm was originally presented for the *minimum* cycle mean problem. However, it is easy to see that the two problems (maximum and minimum) are equivalent, by changing the sign of the edge costs.

Example 3.2: The goal of our second example is to show that the method of finding a "common quadratic Lyapunov function (CQLF)", see e.g. [5], which is one of the most widely used techniques for stability analysis of switched linear systems (or equivalently approximation of the joint spectral radius) is not necessarily exact on rank one matrices. Consider the set of matrices $\mathcal{A} = \{A_1, A_2\}$, with

$$A_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \; A_2 = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}.$$

It is easy to show (e.g. using the maximum cycle mean algorithm) that $\rho(\mathcal{A}) = 1$. However, the CQLF approach can only prove that $\rho(\mathcal{A}) \leq \sqrt{2}$.

References

[1] G. C. Rota and W. G. Strang. A note on the joint spectral radius. *Indag. Math.*, 22:379–381, 1960.

[2] R. Jungers. *The joint spectral radius: theory and applications*, volume 385 of *Lecture Notes in Control and Information Sciences*. Springer, 2009.

[3] V. D. Blondel and J. N. Tsitsiklis. The boundedness of all products of a pair of matrices is undecidable. *Systems and Control Letters*, 41:135–140, 2000.

[4] J. N. Tsitsiklis and V.D. Blondel. The Lyapunov exponent and joint spectral radius of pairs of matrices are hard- when not impossible- to compute and to approximate. *Mathematics of Control, Signals, and Systems*, 10:31–40, 1997.

[5] V. D. Blondel, Yu. Nesterov, and J. Theys. On the accuracy of the ellipsoidal norm approximation of the joint spectral radius. *Linear Algebra Appl.*, 394:91–107, 2005.

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

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

[8] A. A. Ahmadi, R. Jungers, P. A. Parrilo, and M. Roozbehani. Analysis of the joint spectral radius via Lyapunov functions on path-complete graphs. In *Hybrid Systems: Computation and Control 2011*, Lecture Notes in Computer Science. Springer, 2011.

[9] J. Liu and M. Xiao. Rank-one approximation of joint spectral radius of finite matrix family, 2011. `arXiv:1109.1356`.

[10] R.M. Karp. A characterization of the minimum cycle mean in a digraph. *Discrete mathematics*, 23(3):309–311, 1978.

[11] R.K. Ahuja, T.L. Magnanti, and J.B. Orlin. *Network flows: theory, algorithms and applications*. Prentice Hall, 1993.

[12] A. Dasdan and R.K. Gupta. Faster maximum and minimum mean cycle algorithms for system-performance analysis. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 17(10):889–899, 1998.

[13] L. Gurvits and A. Samorodnitsky. A note on common quadratic Lyapunov functions for linear inclusions: exact results and open problems. In *Proceedings of the 44th IEEE Conference on Decision and Control*, 2005.

[14] D. P. Ellerman. Arbitrage theory: a mathematical introduction. *SIAM Review*, 26(2):241–261, 1984.

[15] N. M. Tran. Pairwise ranking: choice of method can produce arbitrarily different rank order, 2011. `arXiv:1103.1110`.

Appendix: Added Remarks for the Revision

After our original submission of this work, we notified the authors of [9] of the error in their characterization of the joint spectral radius of sets of rank one matrices. They have since corrected the mistake and provided an alternative proof of our Corollary 2.1 (see Theorem 3 in Version 6 of [9]), though an algorithm for the computation of the JSR with running time polynomial in n and m is not present in that work.

Also after our initial submission, Vincent Blondel brought to our attention an earlier and independent work of Gurvits and Samorodnitsky [13], where they state (using a slightly different terminology) that the JSR of a set of rank one matrices is given by the maximum cycle gain of $G_{\mathcal{A}}$ [13, Appendix A]. Their work also demonstrates a different and very interesting characterization of the JSR: Given a set of rank one matrices $\mathcal{A} = \{x_1 y_1^T, \ldots, x_m y_m^T\}$, the JSR is less than a positive number a if and only if there exist positive real numbers $d_1, \ldots, d_m$ such that

$$\frac{d_j}{d_i} |y_i^T x_j| \leq a,$$

for all $1 \leq i, j \leq m$. We find it interesting to point out an apparent connection between this characterization and the well-known multiplicative arbitrage theorem in mathematical economics. The theorem states that the exchange rates (or relative prices) between a set of commodities are arbitrage-free if and only if there exist a set of absolute prices for the commodities such that the exchange rates are price ratios [14]. If we treat the entries of the adjacency matrix of $G_{\mathcal{A}}$ as exchange rates, then the arbitrage-free requirement implies that the JSR of the set $\mathcal{A}$ is equal to one. Moreover, the prices promised by the multiplicative arbitrage theorem provide the scalings $d_1, \ldots, d_m$ in the characterization of Gurvits and Samorodnitsky.

Finally, we are grateful to Bernd Sturmfels and Ngoc Tran for pointing out a curious connection between Theorem 2.2 and *tropical eigenvalues*. The theorem in fact shows that the JSR of a set of rank one matrices is equal to the max-plus tropical eigenvalue of the adjacency matrix of $\tilde{G}_{\mathcal{A}}$; see e.g. [15] for a definition. A further exploration of this relationship is left for future research.