Graduate J. Math. 9, Issue 2 (2024), 10 – 14

New proofs for a bound on the spectral radius of the Hadamard geometric mean

Jacob Adamczyk

Abstract

In this short note, we present two new proofs of an inequality first derived by Elsner, Johnson, and Dias Da Silva for an upper bound on the Perron root of the geometric mean of non-negative irreducible matrices. The first proof technique uses the Collatz-Wielandt characterization of the spectral radius; the second establishes a new bound at the matrix element level and a sub-multiplicative property of matrix norms, both of which easily follow from Hölder's inequality.

Keywords: Hadamard geometric mean, eigenvalue inequality, Hölder's inequality.

MSC 2020. Primary: 15A42, Secondary: 15A45

1 Introduction

Let A and B be $n \times n$ irreducible non-negative matrices. We will consider the matrix formed by taking an element-wise geometric mean thereof. To distinguish the standard matrix product from (Hadamard) element-wise operations (product and exponential), we will employ the notation $(\circ)$ as in the following definition.

Definition 1 (Hadamard Geometric Mean of Matrices)**.**

$$C_{ij} \doteq A_{ij}^{\alpha} B_{ij}^{1-\alpha} \doteq \left[A^{(\alpha)} \circ B^{(1-\alpha)} \right]_{ij} \tag{1.1}$$

where the positive real values $\alpha, 1-\alpha$ are the geometric weights.

In this geometric mean setting, [1] provides the following result on the spectral radius of interest, $\rho(C)$:

Theorem 1 (Elsner, Johnson, Dias Da Silva [1])**.** *Let irreducible non-negative matrices $\{A_k\}_{k=1}^{M}$ with corresponding spectral radii $\{\rho(A_k)\}_{k=1}^{M}$ be given, with non-negative weights α_k satisfying $\sum_k \alpha_k \geq 1$. Let $C = A_1^{(\alpha_1)} \circ A_2^{(\alpha_2)} \cdots \circ A_k^{(\alpha_k)}$. Then the spectral radius $\rho(C)$ satisfies:*

$$\rho(C) \leq \prod_{k=1}^{M} \rho(A_k)^{\alpha_k} . \tag{1.2}$$

In the sequel, we consider for simplicity the case of $M = 2$, with the more general case following by induction as in [1]. For convenience, we recall the Collatz-Wielandt formula, which gives a (double-sided) bound on the spectral radius of an irreducible non-negative matrix A, given any choice of a positive vector, $x_i > 0$:

$$\min_i \frac{(Ax)_i}{x_i} \leq \rho(A) \leq \max_i \frac{(Ax)_i}{x_i}. \tag{1.3}$$

Hölder's inequality will also be of importance:

Proposition 1 (Hölder's Inequality)**.** *Let vectors in the non-negative orthant, $\vec{x}, \vec{y} \in \mathbb{R}^N_{\geq 0}$ and exponent $\alpha \in [0,1]$ be given. Then,*

$$\sum_{i=1}^{N} x_i^\alpha y_i^{1-\alpha} \leq \left(\sum_{i=1}^{N} x_i\right)^\alpha \left(\sum_{i=1}^{N} y_i\right)^{1-\alpha}. \tag{1.4}$$

2 First Proof

Our new proof of Theorem 1 (focusing on the case of two matrices with convex geometric weights) now follows straightforwardly from Equations (1.3) and (1.4):

Proof. An upper bound on $\rho(C)$ is given by substituting a variational guess in the right-hand side of Equation (1.3). We choose the geometric mean of constituent right eigenvectors: $x_i = v_i^\alpha w_i^{1-\alpha}$, where $Av = \rho(A)v$ and $Bw = \rho(B)w$. Upon substitution and using Hölder's inequality, this bound simplifies as:

$$\rho(C) \leq \max_i \frac{\sum_{j=1}^{n} (A_{ij}v_j)^\alpha (B_{ij}w_j)^{1-\alpha}}{v_i^\alpha w_i^{1-\alpha}} \leq \max_i \frac{\left(\sum_{j=1}^{n} A_{ij}v_j\right)^\alpha \left(\sum_{j=1}^{n} B_{ij}w_j\right)^{1-\alpha}}{v_i^\alpha w_i^{1-\alpha}} = \rho(A)^\alpha \rho(B)^{1-\alpha} \tag{2.1}$$

□

Although the construction of the variational guess assumed knowledge of the (dominant) right eigenvectors of A and B, the resulting bound depends only on the eigenvalues, a great simplification which seems particular to the geometric mean and upper bound. Of course, if such eigenvectors were indeed known, then one can easily generate a lower bound on $\rho(C)$ via the left-hand side of Equation (1.3). The reverse Hölder's inequality holds [7] for exponents satisfying $\sum_k \beta_k = 1$, such that all $\beta_1, \beta_2, \cdots \beta_{M-1} < 0$ save one, say $\beta_M > 0$. In this setting, Theorem 1 can be trivially extended to such cases by switching the direction of all inequalities in the previous proof.

3 Second Proof

We will now provide an independent proof of Theorem 1, relying on a case of Gelfand's formula:

Proposition 2 (Gelfand's Formula with Frobenius norm)**.** *Let A be an $n \times n$ non-negative matrix. Then, the spectral radius, $\rho(A)$, can be calculated from the Frobenius matrix norm of A as follows:*

$$\rho(A) = \lim_{N\to\infty} \left\| A^N \right\|_F^{1/N}, \tag{3.1}$$

where the Frobenius norm is given by $\|A\|_F^2 = \sum_{i,j} (A_{ij})^2$.

In this section we first prove two sub-multiplicative properties on the level of matrix powers and matrix norms for the Hadamard geometric mean. We then use Gelfand's formula for the spectral radius to give a new proof for Theorem 1, resulting from the more general element-wise inequality in our Theorem 2:

Theorem 2. *Given non-negative $n \times n$ matrices A, B and their Hadamard geometric mean C defined by convex weights $\alpha, 1-\alpha$, the following element-wise inequality holds for all N:*

$$\left(A^{(\alpha)} \circ B^{(1-\alpha)}\right)_{ij}^{N} \doteq C_{ij}^{N} \leq \left(A_{ij}^{N}\right)^{\alpha} \left(B_{ij}^{N}\right)^{1-\alpha}, \tag{3.2}$$

where the power N denotes the standard matrix power.

Proof. We proceed with induction. First, by Equation (1.1) above, the base case ($N = 1$) holds with equality. We assume Equation (3.2) then holds for some $N > 1$ as the inductive hypothesis.

We wish to show

$$C_{ij}^{N+1} \leq \left(A_{ij}^{N+1}\right)^{\alpha} \left(B_{ij}^{N+1}\right)^{1-\alpha}.$$

Writing out the matrix multiplication for C and using the inductive hypothesis gives:

$$C_{ij}^{N+1} = \sum_{l=1}^{n} C_{il}^{N} C_{lj} \leq \sum_{l=1}^{n} \left(A_{il}^{N}\right)^{\alpha} \left(B_{il}^{N}\right)^{1-\alpha} C_{lj} = \sum_{l=1}^{n} \left(A_{il}^{N} A_{lj}\right)^{\alpha} \left(B_{il}^{N} B_{lj}\right)^{1-\alpha}.$$

Recall Proposition 1 and consider a fixed i, j. Then viewing the factors in the sum as two vectors in $\mathbb{R}^n$, we can directly apply Hölder's inequality to the right-hand side, giving:

$$C_{ij}^{N+1} \leq \left[\sum_{l=1}^{n} A_{il}^{N} A_{lj}\right]^{\alpha} \left[\sum_{l=1}^{n} B_{il}^{N} B_{lj}\right]^{1-\alpha} = \left(A_{ij}^{N+1}\right)^{\alpha} \left(B_{ij}^{N+1}\right)^{1-\alpha}, \tag{3.3}$$

which concludes the inductive proof. □

Theorem 1 can now be proven via our Theorem 2 after making use of the following property of the Frobenius matrix norm for the Hadamard geometric mean of matrices. The following sub-multiplicative property holds for the standard matrix product and a variety of other matrix operations, cf. [6].

Lemma 1. *Let non-negative $n \times n$ matrices A, B and $\alpha \in [0, 1]$ be given. The Frobenius matrix norm has the following sub-multiplicative property for the Hadamard geometric means of matrices:*

$$\left\|A^{(\alpha)} \circ B^{(1-\alpha)}\right\|_F \leq \|A\|_F^{\alpha} \cdot \|B\|_F^{1-\alpha}. \tag{3.4}$$

Proof. Writing out the definition of the Frobenius norm and the Hadamard geometric mean,

$$\left\|A^{(\alpha)} \circ B^{(1-\alpha)}\right\|_F^2 = \sum_{i,j=1}^{n} \left(A_{ij}\right)^{2\alpha} \left(B_{ij}\right)^{2(1-\alpha)}. \tag{3.5}$$

We now interpret the squared matrix elements $(A_{ij})^2$ and $(B_{ij})^2$ as two vectors in $\mathbb{R}^{n^2}$, and apply Hölder's inequality,

$$\left\|A^{(\alpha)} \circ B^{(1-\alpha)}\right\|_F^2 \leq \left(\sum_{i,j=1}^{n} \left(A_{ij}\right)^2\right)^{\alpha} \left(\sum_{i,j=1}^{n} \left(B_{ij}\right)^2\right)^{1-\alpha} = \left(\|A\|_F^2\right)^{\alpha} \cdot \left(\|B\|_F^2\right)^{1-\alpha}, \tag{3.6}$$

which gives the required result after taking a square root. □

Remark. *Lemma 1 holds for the more general case of the $L_{p,q}$ matrix norm, but the special case $p = q = 2$ shown above is sufficient for our purposes.*

Below we provide a new proof of Theorem 1 in light of the preceding results.

Proof. Starting from Gelfand's formula for the spectral radius (Equation (3.1)), and applying Theorem 2 and Lemma 1 in succession:

$$\begin{aligned}\rho(C) &= \lim_{N\to\infty} \left\|C^N\right\|_F^{1/N} \le \lim_{N\to\infty} \left\|\left(A^N\right)^{(\alpha)} \left(B^N\right)^{(1-\alpha)}\right\|_F^{1/N} \\ &\le \lim_{N\to\infty} \left\|A^N\right\|_F^{\alpha/N} \left\|B^N\right\|_F^{(1-\alpha)/N} = \left(\lim_{N\to\infty} \left\|A^N\right\|_F^{1/N}\right)^{\alpha} \left(\lim_{N\to\infty} \left\|B^N\right\|_F^{1/N}\right)^{1-\alpha} \\ &= \rho(A)^{\alpha}\rho(B)^{1-\alpha}.\end{aligned}$$

□

4 Discussion

In this short note, we provided two new proofs of the upper bound on the spectral radius for the Hadamard geometric mean of non-negative matrices. The original proofs in [1] used the log-convexity of matrix elements and a construction based on the maximum element of the eigenvector corresponding to $\rho(C)$. On the other hand, the proofs presented here are arguably more transparent: requiring no auxiliary constructions and instead using the definitions of the Perron root directly. Although our proofs considered the case of two matrices to construct the geometric mean, all results extend trivially to the case of $M > 2$ matrices and super-convex weights, as noted.

In average-reward reinforcement learning, the reward rate and differential value function are related to the logarithm of the Perron root and associated eigenvector of a certain "tilted transition matrix" [4], respectively. Based on this connection, inductive proofs used for the non-linear Bellman equation [5] suggest the application of similar approaches to study the Perron root (as discussed in Section 3). These techniques build on the analogy between geometric means of matrices and convex combinations of rewards (with the last row of Table 2 in [5] corresponding to the logarithm of our Eq. (1.2)). Based on the remaining connection to the differential value function, it may be possible to find similar results concerning the corresponding Perron eigenvectors.

In future work, it may be of interest to extend such proofs to multilinear forms. It may also be possible to consider reversed Hölder's inequalities, such as in [3], alongside the Collatz-Wielandt formulation to obtain eigenvector-independent double-sided bounds.

5 Acknowledgements

The author expresses his gratitude to Rahul V. Kulkarni, Alexander Moll, and Mirjana Vuletić for helpful discussions. He is thankful for the financial support provided from the NSF through Award No. DMS-1854350.

References

[1] L. Elsner and Charles R. Johnson and J.A. Dias Da Silva, textitThe Perron root of a weighted geometric mean of nonneagative matrices, Linear and Multilinear Algebra, **24** (1988), 1–13.

[2] Edwin F. Beckenbach and Richard Bellman, *Inequalities*, Springer Berlin Heidelberg (1961).

[3] Zhao, Chang-Jian and Cheung, Wing Sum, *Hölder's reverse inequality and its applications*, Publications de l'Institut Mathematique, **99** (2016), 211–216.

[4] Arriojas, Argenis and Adamczyk, Jacob and Tiomkin, Stas and Kulkarni, Rahul V, *Entropy regularized reinforcement learning using large deviation theory*, Physical Review Research, **5** (2023), 023085.

[5] Adamczyk J., Makarenko V., Arriojas A., Tiomkin S., Kulkarni R. V., *Bounding the optimal value function in compositional reinforcement learning*, Uncertainty in Artificial Intelligence, PMLR, (2023), 22–32.

[6] Horn R. A., Johnson C. R., *Matrix analysis*, Cambridge University Press (2012).

[7] Beckenbach E. F., *On Hölder's inequality*, J. Math. Anal. Appl., **15** (1966), 21–29.

Jacob Adamczyk

Department of Physics,
University of Massachusetts Boston,
100 Morrissey Blvd., Boston, 02125, MA, USA

E-mail address: jacob.adamczyk001@umb.edu