Communication

Jeffrey Uhlmann*

Class of finite-dimensional matrices with diagonals that majorize their spectrum

https://doi.org/10.1515/spma-2022-0185
received October 30, 2022; accepted January 06, 2023

Abstract: We define a special class of finite-dimensional matrices for which the diagonal majorizes the spectrum. This is the first class of matrices known to have this property, although the reverse majorization (i.e., the spectrum majorizing the diagonal) was previously known to hold for unitarily diagonalizable (i.e., normal) matrices. Currently, these are the only known matrix classes that structurally provide a majorization relationship between their spectrum and diagonal.

Keywords: Inverse relative gain array, PD diagonalizable, PD-IRGA, majorization, positive-definite similarity, special matrices

MSC 2020: 15-06, 15A57

1 Introduction

Majorization provides a preorder relational operator for comparing the relative disorder (entropy) of the distribution of elements of two vectors. Specifically, given two vectors $x, y \in \mathbb{R}^n$, whose respective elements sum to unity[1], then y is said to *majorize* x, notated as $x \prec y$, if and only if there exists a set of permutation matrices P_j and probabilities p_j, such that

$$x = \sum_j p_j P_j y. \tag{1}$$

This expression of x as a probabilistic sum of permutations of y can be interpreted as indicating that the distribution of elements of x is more *disordered* than that of y. An alternative definition is that $x \prec y$ if and only if

$$x = Sy, \tag{2}$$

where S is a doubly stochastic matrix, i.e., nonnegative with every row and column sum equal to unity. This definition makes the probabilistic interpretation even more explicit in the form $Sy \to x$, where x is seen to be a conservative stochastic evolutionary state of y. This can be made rigorous by the following known result:

$$x \prec y \;\Rightarrow\; H(x) \geq H(y), \tag{3}$$

where $H(\cdot)$ is the Shannon entropy, $H(z) \doteq -\sum_{i=1}^{n} z_i \log z_i$, where $z_i \log z_i$ is taken as 0 for $z_i = 0$.

1 Summing to unity is needed only for strict probabilistic interpretations but is not generally required to define majorization.

* **Corresponding author: Jeffrey Uhlmann,** Department of Electrical Engineering and Computer Science, University of Missouri-Columbia, 201 Naka Hall, Columbia, MO 65211, United States, e-mail: uhlmannj@missouri.edu

Open Access. © 2023 the author(s), published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.

The Schur-Horn theorem [2,8] says that every Hermitian matrix has the property that its spectrum (vector of eigenvalues) majorizes its diagonal, and the same can be inferred more generally from equation (2) for all normal, i.e., unitarily diagonalizable, matrices. This structural property immediately provides an additional means for analyzing, modeling, and controlling systems involving such matrices. A prominent example from physics is the density matrix describing the quantum state of a system, which is Hermitian (more specifically, positive semidefinite) and therefore normal.

In the case of a normal matrix, the majorization property implies that its diagonal provides an entropic upper bound on the distribution of its eigenvalues. Having such a link between the diagonal and spectrum of a given matrix offers a variety of potential theoretical and practical benefits, but heretofore there has been no other matrix class for which such a majorization property has been proven to hold.

In Section 2, we identify a new finite-dimensional matrix class for which an alternative diagonal/spectrum majorization relationship can be proven to exist.

2 Special PD-diagonalizable matrices

We begin by stating a previously studied [13] conjecture/result:

PD-IRGA conjecture: Given any real $n \times n$ symmetric positive-definite (PD) matrix P, then for $n \leq 6$:

$$S = (P \circ P^{-1})^{-1} \in \text{ nonsingular nonnegative PD doubly stochastic}, \tag{4}$$

where "$\circ$" is the Schur-Hadamard elementwise matrix product.

Such a result would be remarkable because it implies that important properties exist for low-dimensional PD matrices that do not hold generally. Specifically, it can be shown by counterexample that nonnegativity does not generally hold for $n > 6$, and it has been proven for all $n \leq 4$ (see ref. [13] for details).

Finite-dimensional properties of this kind are of particular interest when analyzing systems that are intrinsically low dimensional due to constraints imposed by physical theory, e.g., the three spatial dimensions of the physical world. Here, however, we focus on the appearance of the doubly stochastic matrix S and observe that it can permit us to define a special finite-dimensional matrix class for which the diagonal majorizes the spectrum. Specifically, letting k be the largest integer for which the PD-IRGA conjecture holds, we define the following matrix form:

Definition. (Special PD diagonalizable) A given $n \times n$ matrix $M = \mathrm{PEP}^{-1}$, $n \leq k$, is defined to be special PD diagonalizable if matrix E is diagonal and P is symmetric (real) PD.

The motivation for this definition is the following:

Theorem 1. *The spectrum of a special PD diagonalizable matrix is majorized by its diagonal.*

The proof relies on the known result [9] that for a diagonalizable matrix $B = AEA^{-1}$:

$$(A \circ A^{-T}) \cdot \operatorname{diag}(E) = \operatorname{diag}(B), \tag{5}$$

where A^{-T} is the transposed[2] inverse of A. In other words, the matrix $A \circ A^{-T}$ – *which may be singular even if A is not* – maps the spectrum of B to the diagonal of B. If A is real symmetric PD, then $A \circ A^{-1}$ is nonsingular, so multiplying both sides by its inverse $S = (A \circ A^{-1})^{-1}$ gives

$$\operatorname{diag}(E) = S \cdot \operatorname{diag}(B), \tag{6}$$

2 Note that the transpose operator (not the complex conjugate) applies even in the case of complex A.

where S is doubly stochastic. Thus, by the majorization criterion of equation (2), it can be inferred that the diagonal of a special PD-diagonalizable matrix B majorizes its spectrum:

$$\mathrm{diag}(E) \prec \mathrm{diag}(B), \tag{7}$$

and, therefore, its entropy gives a lower bound on that of its spectrum, i.e.,

$$H(\mathrm{diag}(E)) \geq H(\mathrm{diag}(B)). \tag{8}$$

3 Potential applications

General PD-similarity transformations appear in a variety of theoretical and practical applications. These include Gaussian orthogonal ensembles from random matrix theory [10] and their applications to pseudo/quasi-Hermitian operators in physics [5,12], e.g., Su-Schrieffer-Heeger models with real eigenvalues, for which a non-Hermitian Hamiltonian H is mapped to a Hermitian Hamiltonian $\hat{H}$ by a PD similarity transformation [3,4,11]. It can be observed that for real[3] H and $\hat{H}$, a real orthogonal transformation of the system that diagonalizes $\hat{H}$ will transform H to PD-diagonalizable form. This can be shown by expressing the normal matrix $\hat{H}$ as follows:

$$\hat{H} = UEU^T \tag{9}$$

and its PD-similar companion H as follows:

$$H = P\hat{H}P^{-1}, \tag{10}$$

$$= (VDV^T)\,\hat{H}(VDV^T)^{-1}. \tag{11}$$

Applying the diagonalizing similarity transformation U to H gives

$$U^T HU = U^T(VDV^T)\hat{H}(VDV^T)^{-1}U, \tag{12}$$

$$= U^T(VDV^T)(UEU^T)(VDV^T)^{-1}, U, \tag{13}$$

$$= (U^T VDV^T U)E(U^T VDV^T U)^{-1}, \tag{14}$$

$$= ((U^T V)D(U^T V)^T)E\ ((U^T V)D(U^T V)^T)^{-1}, \tag{15}$$

$$= PEP^{-1}, \tag{16}$$

where $P = (U^T V)D(U^T V)^T$. If $n \leq k$, then $U^T HU$ is special PD diagonalizable and, therefore, its diagonal majorizes its spectrum. In other words, under these conditions, the matrix H can be orthogonally transformed so that its diagonal majorizes the spectrum of the potentially unknown Hamiltonian $\hat{H}$. Alternatively, the process can be viewed from an optimization perspective as a problem of identifying an orthogonal-similarity transformation W such that the Shannon entropy of diag(WHW^T) is optimized via majorization to provide a best-possible provable bound on the entropy of its spectrum.

The PD-IRGA conjecture originated in the area of control theory [13], and special PD-diagonalizable matrices can be examined in this context with the spectrum representing the state of a system. It should be noted that the intrinsic representable dimensionality of this state is not necessarily limited to $k \leq 6$ when eigenvalues provide additional degrees of freedom. Specifically, while the doubly stochastic mapping from spectrum to diagonal in equation (6) requires the diagonalizing matrix to be real PD, that condition is not required for the spectrum if majorization is generalized (see ref. [2], p. 110) from vectors over $\mathbb{R}^n$ to $\mathbb{C}^n$, and more generally to their normed division-algebra extensions over $\mathbb{H}^n$ and $\mathbb{O}^n$. Thus, if the spectrum is defined

3 Subsequent discussion of generalizations of majorization to complex vectors (and other cases) can be implicitly applied to partially relax the restriction to reals in this example. Further generalizations are discussed in the Appendix.

over the latter, i.e., octonions, then the maximum representable dimensionality is $8 \cdot k$, which gives 48 degrees of freedom for the conjectured value of $k = 6$.

Generalizations of majorization to state vectors over hypercomplex spaces (or other algebras, e.g., $*$-algebras) have important implications for practical applications when computation of the spectrum is prohibitively expensive. For example, consider a measurement process that is designed to obtain a special PD-diagonalizable matrix X and normal matrix Y, both having the same spectrum, i.e., the state distribution of interest. If it is not practical to compute the spectrum explicitly, then Theorem 1 implies that lower and upper entropic bounds on the measured state can be obtained from the diagonals of X and Y, respectively. These bounds may be sufficient for the needs of the application or, if not, may provide a means for guiding a search of the space toward the case in which the measurement matrices become equal, i.e., diagonal and give the exact spectrum of the target state.

4 Discussion

The principal contribution of this article is the definition of a new class of finite-dimensional matrices, *special PD-diagonalizable*, for which there exists a majorization relationship between the diagonal and the spectrum. The definition of this class exploits the proven cases of the PD-IRGA conjecture, which presently is restricted to a subset of $n \times n$ matrices with $n \leq 4$, but is conjectured to hold for $n \leq 6$.

Future work will examine applications of these results to tracking and control problems in which measurements of an evolving state are transformed by PD similarity with respect to the covariant derivative along a Riemannian manifold [6,7], which may also be applicable to models arising in both theoretical and applied physics. Of particular focus for the latter will be potential applications to robust quantum circuits [14]. From a theoretical physics perspective, understanding the structural constraints imposed by dimensionality $d \leq 4$ or $d \leq 6$ that guarantee the properties of special PD-diagonalizable matrices may reveal deeper insights relating to quantum field theories that are strictly constrained to dimensions $d = 4$ or $d = 6$. (e.g., conformal field theory [15] and loop quantum gravity [1]).

Conflict of interest: The author reports no conflicting interests or funding.

Data availability statement: Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

References

[1] A. Ashtekar and E. Bianchi, *A short review of loop quantum gravity*, Rep. Prog. Phys. **84** (2021), 042001.
[2] R. B. Bapat and V. S. Sunder, *On majorization and Schur products*, Linear Algebra Appl. **72** (1985), 107–117.
[3] C. M. Bender and S. Boettcher, *Real spectra in non-Hermitian hamiltonians having PT symmetry*, Phys. Rev. Lett. **80** (1998), 5243.
[4] M. Bertolotti, *PT symmetry in quantum and classical physics*, Contemporary Phys. **60** (2019), 2.
[5] J. Dieudonné, *Quasi-Hermitian operators*, in: Proceedings Of the International Symposium on Linear Spaces, 1961 July, pp. 115–123.
[6] M. J. Eilders, *Decentralized Riemannian particle filtering with applications to multi-agentlocalization*, Air Force Institute of Technology, AFIT Scholar Theses and Dissertations, 2012, https://scholar.afit.edu/etd/1101.
[7] S. Hauberg, F. Lauze, and K. S. Pederson, *Unscented Kalman filtering on Riemannian manifolds*, J. Math. Imaging Vision, **46** (2013), 103–120.
[8] A. Horn, *Doubly stochastic matrices and the diagonal of a rotation matrix*, Amer. J. Math. **76** (1954), 620–630.
[9] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.
[10] M. L. Mehta, *Random Matrices*, Elsevier, Amsterdam, 2004.

[11] K. Mochizuki, N. Hatano, J. Feinberg, and H. Obuse, *Statistical properties of eigenvalues of the non-Hermitian Su-Schrieffer-Heeger model with random hopping terms*, Phys. Rev. E **102** (2020), 012101.
[12] A. Mostafazadeh, *Pseudo-supersymmetric quantum mechanics and isospectral pseudo-Hermitian Hamiltonians*, Nuclear Phys. B **640** (2002), no. 3, 419–434.
[13] J. Uhlmann and J. Wang, *On radically expanding the landscape of potential applications for automated-proof methods*, SN Computer Science, (Springer-Nature), **2** (2021), 259.
[14] H. Wang, Y. Ding, J. Gu, Z. Li, Y. Lin, D. Z. Pan, et al., *QuantumNAS: Noise-adaptive search for robust quantum circuits*, in: The 28th IEEE International Symposium on High-Performance Computer Architecture (HPCA), IEEE, Seoul, Korea, 2022.
[15] E. Witten, *Conformal Field Theory in Four and Six Dimensions*, 2007, arXiv:0712.0157.

Appendix

The PD-IRGA conjecture/theorem actually applies more generally to real PD matrices P that are transformed on the left and right by arbitrary permutations and/or nonsingular diagonal matrices. In other words, the structure of the PD matrix P can be generalized to the form:

$$X = D_L Q_L P Q_R D_R,$$

where Q_L and Q_R are permutation matrices, P is real PD, and D_L and D_R are nonsingular diagonal. The proof follows directly by verifying that $A \circ A^{-T}$ from equation (5) (the transpose was unnecessary in the original PD-IRGA expression because P was symmetric) is invariant with respect to nonsingular diagonal transformations of the argument A and equivariant with respect to permutations of A.

Thus, the properties defining a special PD-diagonalizable matrix $M = PEP^{-1}$ apply also to $M = XEX^{-1}$. Thus, Theorem 1 generalizes to guarantee that the diagonal of a matrix M of the form

$$M = (D_L Q_L P Q_R D_R) E (D_L Q_L P \ Q_R D_R)^{-1}$$

majorizes its spectrum, i.e., diag(E). There are a multitude of obvious generalizations of the example of equations (9)–(16), but at present, there are no obvious applications for which such generality is likely to be exploited.