AN OPERATOR VERSION OF SINKHORN'S THEOREM IN EUCLIDEAN JORDAN ALGEBRAS

M. SEETHARAMA GOWDA[1,*]

[1]Department of Mathematics and Statistics
University of Maryland Baltimore County
Baltimore, Maryland 21250, USA

Abstract. Sinkhorn's theorem asserts that if A is an $n \times n$ matrix with positive entries, then there are positive diagonal matrices C and D, unique up to scaling factors, such that CAD is a doubly stochastic matrix. We generalize this to Euclidean Jordan algebras: If P is a strictly positive linear map on a Euclidean Jordan algebra $\mathscr{V}$, then there exist positive elements c and d in $\mathscr{V}$, unique up to scaling factors, such that P_cPP_d is a doubly stochastic map on $\mathscr{V}$, where P_c and P_d are quadratic representations of c and d respectively.

Keywords. Sinkhorn's theorem; Euclidean Jordan algebra; strictly positive map; doubly stochastic map; quadratic representation.

2020 Mathematics Subject Classification. 15B51, 15B57, 17C20, 47B65.

Dedication: This paper is dedicated to the memory of Thomas Seidman.

1. INTRODUCTION

Sinkhorn's theorem [12] asserts that if A is an $n \times n$ matrix with positive entries, then there are positive diagonal matrices C and D, unique up to scaling factors, such that CAD is doubly stochastic (meaning all its row/column sums are 1). Since its appearance in 1964, this result has been extended in numerous directions with varying proofs and applications, see [9] for an exhaustive review/survey. With the observation that $\mathscr{R}^n$ is a Euclidean Jordan algebra in which quadratic representations are diagonal matrices, we extend Sinkhorn's result to general Euclidean Jordan algebras. To elaborate, consider a Euclidean Jordan algebra $(\mathscr{V}, \circ, \langle \cdot, \cdot \rangle)$ of rank n with $\circ$ denoting the Jordan product and $\langle \cdot, \cdot \rangle$ denoting the (trace) inner product. Let $\mathscr{V}_+$ denote the corresponding (symmetric) cone of squares with interior $\mathscr{V}_{++}$. A linear map $P : \mathscr{V} \to \mathscr{V}$ is said to be *positive* if $P(\mathscr{V}_+) \subseteq \mathscr{V}_+$ and *strictly positive* if $P(\mathscr{V}_+ \setminus \{0\}) \subseteq \mathscr{V}_{++}$.
A positive linear map P on $\mathscr{V}$ is said to be *doubly stochastic* if it is *unital* and *trace preserving*, meaning that $P(e) = e$ and $P^*(e) = e$, where e denotes the unit element of $\mathscr{V}$ and P^* denotes the adjoint/conjugate of P. Given $a \in \mathscr{V}$, we consider the *quadratic representation* P_a of a defined by

$$P_a(x) := 2a \circ (a \circ x) - a^2 \circ x \quad (x \in \mathscr{V}). \tag{1.1}$$

We now formulate our generalization as follows:

*Corresponding author.
E-mail address: gowda@umbc.edu.
Received xx, x, xxxx; Accepted xx, x, xxxx.

Theorem 1.1. *Suppose P is a strictly positive linear map on $\mathcal{V}$. Then there exist $c,d \in \mathcal{V}_{++}$, unique up to scaling factors, such that $P_c P P_d$ is doubly stochastic.*

That Sinkhorn's result is a special case of this theorem can be seen by considering the Euclidean Jordan algebra $\mathcal{R}^n$ where the Jordan product is the componentwise product and the inner product is the usual one. (In this case, a strictly positive linear map P on $\mathcal{R}^n$ is just a matrix with positive entries, and the quadratic representations P_c and P_d become positive diagonal matrices.) Another special case is obtained by considering the algebra $\mathcal{H}^n$ of all $n \times n$ complex Hermitian matrices with Jordan product $X \circ Y := \frac{XY+YX}{2}$, inner product $\langle X,Y\rangle := tr(XY)$, and (Identity matrix) I as the unit element. In this setting, for any $A \in \mathcal{H}^n$, the quadratic representation of A is given by $P_A(X) := AXA$. So, if $P : \mathcal{H}^n \to \mathcal{H}^n$ is a strictly positive linear map, then there are (Hermitian) positive definite matrices C and D such that the map $X \to CP(DXD)C$ is doubly stochastic on $\mathcal{H}^n$. Results of the latter type appear in the literature on positive maps on the space of all $n \times n$ complex matrices, see e.g., the weaker form of Gurvitz generalization to positive maps (Theorem 1.2 in [9]), Corollary 9.16 in [9], and general results by Gurvitz (Theorems 9.4 and 9.5 in [9]). These are achieved using a variety of concepts/results/tools such as the Brower fixed point theorem, Hilbert metric, capacity, etc. Many of them also deal with computational aspects.

In our approach to proving Theorem 1.1, we use standard arguments: The uniqueness is proved via a convexity argument, while the existence part uses the Brouwer fixed point theorem. Our contribution here is to demonstrate the use of elegant Euclidean Jordan algebra machinery to achieve the above-mentioned generalization. In the setting of the algebra of (complex) Hermitian matrices, a far-reaching generalization of Sinkhorn's theorem due to Gurvitz [7] exists. This generalization is expressed in terms of fully indecomposable maps and capacity. At present, these concepts are not available in the setting of Euclidean Jordan algebras. We anticipate a complete generalization when these are available and developed.

2. Preliminaries

Throughout, $(\mathcal{V}, \circ, \langle\cdot,\cdot\rangle)$ denotes a Euclidean Jordan of rank n with unit element e. We freely use various notations, properties, etc., from the literature [2, 5]. In $\mathcal{V}$, every x has a spectral decomposition: $x = \sum_{i=1}^n x_i e_i$, where the real numbers $x_1, x_2, \ldots, x_n$ are (called) the eigenvalues of x and $\{e_1, e_2, \ldots, e_n\}$ is a Jordan frame. We write

$$\lambda(x) = \Big(\lambda_1(x), \lambda_2(x), \ldots, \lambda_n(x)\Big)$$

for the vector of eigenvalues of x, *written in the decreasing order.* For $x \in \mathcal{V}$, trace of x is defined by $\mathrm{tr}(x) := \sum_{i=1}^n \lambda_i(x)$. Then, the inner product $(x,y) \mapsto tr(x \circ y)$ defines another inner product on $\mathcal{V}$ that is compatible with the given Jordan product. *Henceforth, we will assume that $\mathcal{V}$ carries this (trace) inner product.* (So, elements of a Jordan frame have unit norm and are mutually orthogonal.)

In $\mathcal{V}$, we write $x \geq 0$ when $\lambda(x) \geq 0$ in $\mathcal{R}^n$; likewise $x > 0$ when $\lambda(x) > 0$. We also write $x \geq y$ when $x - y \geq 0$, etc. The symmetric cone of $\mathcal{V}$ is defined as $\mathcal{V}_+ := \{x \in \mathcal{V} : x \geq 0\}$; its interior is $\mathcal{V}_{++} = \{x \in \mathcal{V} : x > 0\}$. Corresponding to the spectral decomposition $x = \sum_{i=1}^n x_i e_i$, we define, whenever appropriate, $\sqrt{x} := \sum_{i=1}^n \sqrt{x_i} e_i$ and $x^{-1} := \sum x_i^{-1} e_i$. We freely use the following implication (which is a consequence of the min-max theorem of Hirzebruch , see [6], Theorem 3.1):

$$x \leq y \Rightarrow \lambda(x) \leq \lambda(y).$$

Recall that a linear map $P : V \to \mathcal{V}$ is positive if $P(\mathcal{V}_+) \subseteq \mathcal{V}_+$ and strictly positive if $P(\mathcal{V}_+ \setminus \{0\}) \subseteq \mathcal{V}_{++}$. As $\mathcal{V}_+$ is a self-dual cone, we have: If P is positive (strictly positive), then so is P^* (respectively, strictly positive). A linear map P is *doubly stochastic* if it is positive, unital (meaning $P(e) = e$) and trace-preserving (that is, $\mathrm{tr}(P(x)) = \mathrm{tr}(x)$ for all $x \in \mathcal{V}$ or, equivalently, $P^*(e) = e$).

For any $a \in \mathcal{V}$, we define the quadratic representation P_a of a by (1.1). Below are some properties of P_a [2] that we frequently use:

- P_a is self-adjoint and positive. Also, $P_a(e) = a^2$.
- When a is invertible (for example, $a > 0$), $(P_a)^{-1} = P_{a^{-1}}$ and $x > 0 \Rightarrow P_a(x) > 0$.
- Whenever defined, $(P_a(x))^{-1} = P_{a^{-1}}(x^{-1})$. In particular, this holds when $a, x > 0$.

The following result characterizes strictly positive maps and also provides an example.

Proposition 2.1. *Let $E : \mathcal{V} \to \mathcal{V}$ be the map defined by $E(x) := tr(x)\,e$. Then, a linear map P is strictly positive on $\mathcal{V}$ if and only if there is a positive number α such that*

$$P(x) \geq \alpha E(x) \quad \textit{for all } x \geq 0.$$

Proof. As the map E is strictly positive, the 'if' part is obvious; we prove the 'only if' part. Suppose P is strictly positive on $\mathcal{V}$, but there is no α with the specified property. Then, for each natural number k, there exists $x_k \geq 0$ with $||x_k|| = 1$ such that

$$P(x_k) - \frac{1}{k}\langle x_k, e\rangle\, e \notin \mathcal{V}_+.$$

Since $\mathcal{V}_+$ is a closed convex cone, by a standard separation theorem, for each k, there exists an element $d_k \in \mathcal{V}$ such that $||d_k|| = 1$ and

$$\Big\langle d_k, P(x_k) - \frac{1}{k}\langle x_k, e\rangle\, e \Big\rangle < 0 \leq \langle d_k, y\rangle \quad \text{for all } y \in \mathcal{V}_+.$$

As $\mathcal{V}_+$ is a self-dual cone, from the inequality on the right, we see that $d_k \in \mathcal{V}_+$. Now, we may assume, without loss of generality, that $x_k \to \overline{x}$ and $d_k \to \overline{d}$. Then, $\overline{x}, \overline{d} \geq 0$ and $||\overline{x}|| = 1 = ||\overline{d}||$. Letting $k \to \infty$ in the inequality $\big\langle d_k, P(x_k) - \frac{1}{k}\langle x_k, e\rangle\, e \big\rangle < 0$, we get

$$\langle \overline{d}, P(\overline{x})\rangle \leq 0.$$

As P is stricly positive and $0 \neq \overline{x} \geq 0$, we have $P(\overline{x}) > 0$. This implies, as $0 \leq \overline{d} \neq 0$, $\langle \overline{d}, P(\overline{x})\rangle > 0$. We reach a contradiction. This proves the 'only if' part. □

Towards the proof of our main theorem, we prove the following lemmas.

Lemma 2.2. *Let P be a positive map and $a = P(e)$. Then, for all $x \geq 0$ and $i = 1, 2, \ldots, n$,*

$$\lambda_n(x)\,\lambda_i(a) \leq \lambda_i(P(x)) \leq \lambda_1(x)\,\lambda_i(a).$$

In particular, when $a = P(e) = e$,

$$\lambda_n(x) \leq \lambda_i(P(x)) \leq \lambda_1(x) \quad \textit{for all } x \geq 0,\ 1 \leq i \leq n. \tag{2.1}$$

Proof. Fix $x \geq 0$. Then, we have $x - \lambda_n(x)\, e \geq 0$ and $\lambda_1(x)\, e - x \geq 0$. Hence, $P(x) - \lambda_n(x) P(e) \geq 0$ and $\lambda_1(x) P(e) - P(x) \geq 0$. This gives

$$\lambda_n(x)\, a \leq P(x) \leq \lambda_1(x)\, a.$$

Using the implication $u \leq v \Rightarrow \lambda(u) \leq \lambda(v)$ and noting that $x \geq 0$, we get

$$\lambda_n(x)\, \lambda(a) \leq \lambda(P(x)) \leq \lambda_1(x)\, \lambda(a).$$

In particular, when $a = P(e) = e$, we have $\lambda_i(a) = \lambda_i(e) = 1$ for all i and (2.1) follows. $\square$

Lemma 2.3. *Let P be strictly positive and doubly stochastic on $\mathcal{V}$. Suppose $u, v > 0$ with $P(u) = v^{-1}$ and $P^*(v) = u^{-1}$. Then, there exists a (scalar) $t > 0$ such that $u = t\, e$ and $v = \frac{1}{t}\, e$.*

Proof. We write the spectral decompositions $u = \sum_1^n u_i e_i$ and $v = \sum_1^n v_i f_i$, where, without loss of generality, $u_1 \geq u_2 \geq \cdots \geq u_n > 0$ and $v_1 \geq v_2 \geq \cdots \geq v_n > 0$. Then, $u^{-1} = \sum_1^n u_i^{-1} e_i$ and $v^{-1} = \sum_1^n v_i^{-1} f_i$. As P is doubly stochastic, $P(e) = e = P^*(e)$. Now, from (2.1) with $x = u$ and $P(x) = P(u) = v^{-1}$, we get

$$u_n \leq \frac{1}{v_i} \leq u_1 \quad (1 \leq i \leq n).$$

Again, applying (2.1) to P^* with $x = v$ and $P^*(x) = P^*(v) = u^{-1}$, we get

$$v_n \leq \frac{1}{u_i} \leq v_1 \quad (1 \leq i \leq n).$$

From these inequalities, we see that

$$\frac{1}{u_n} = v_1 \quad \text{and} \quad \frac{1}{v_n} = u_1.$$

Now, $v^{-1} = P(u) = P(\sum_1^n u_i e_i)$ implies

$$\sum_{i=1}^n \frac{1}{v_i} f_i = P(u) = \sum_{j=1}^n u_j P(e_j).$$

Taking the inner product of the above quantity with f_i (for any $i = 1, 2, \ldots, n$), we get

$$\frac{1}{v_i} = \sum_{j=1}^n \langle f_i, P(e_j) \rangle u_j.$$

As P is doubly stochastic, it is easy to see that the matrix $A = [a_{ij}]$ with $a_{ij} = \langle f_i, P(e_j) \rangle$ is a doubly stochastic matrix (see [4], Theorem 5), that is, the entries of A are nonnegative and each row/column sum in A is 1. Now, from above,

$$\frac{1}{v_n} = \sum_{j=1}^n \langle f_n, P(e_j) \rangle u_j = \sum_{j=1}^n a_{nj} u_j \leq \Big(\sum_{j=1}^n a_{nj} \Big) u_1 = u_1,$$

as u_1 is the largest of all u_js. Since $\frac{1}{v_n} = u_1$ (proved earlier), the above inequality turns into an equality. As P is strictly positive, $P(e_j) > 0$ and so $\langle f_i, P(e_j) \rangle > 0$ for all i, j. Thus, $a_{nj} > 0$ for all j. As $u_1 \geq u_2 \geq \cdots \geq u_n$, the equality $\frac{1}{v_n} = \sum_1^n a_{nj} u_j = u_1$ shows that $u_1 = u_2 = \cdots = u_n$, that is, $u = \sum_1^n u_i e_i = u_1 (\sum_1^n e_i) = u_1 e$. Putting $t = u_1$, we see that $t > 0$ and $u = te$. Then $v^{-1} = P(te) = tP(e) = te$ implies that $v = \frac{1}{t} e$. $\square$

3. Proof of Theorem 1.1

We first prove the uniqueness part. Suppose $a,b,c,d>0$ such that $T:=P_aPP_b$ and $S:=P_cPP_d$ are doubly stochastic on $\mathcal{V}$. Using the relations $(P_a)^{-1}=P_{a^{-1}}$, etc., we see that

$$T=P_aP_{c^{-1}}SP_{d^{-1}}P_b.$$

Then,

$$T(e)=e\Rightarrow P_aP_{c^{-1}}SP_{d^{-1}}P_b(e)=e\Rightarrow SP_{d^{-1}}P_b(e)=\Big(P_aP_{c^{-1}}\Big)^{-1}(e)=P_cP_{a^{-1}}(e)=P_c(a^{-2}).$$

Letting

$$u:=P_{d^{-1}}P_b(e)=P_{d^{-1}}(b^2)\quad\text{and}\quad v:=P_{c^{-1}}P_a(e)=P_{c^{-1}}(a^2),\tag{3.1}$$

we see that $S(u)=v^{-1}$. Similarly, $T^*(e)=e$ results in $S^*(v)=u^{-1}$. Since $c,d>0$ and P is strictly positive, we have

$$0\neq x\geq 0\Rightarrow 0\neq P_d(x)\geq 0\Rightarrow P(P_d(x))>0\Rightarrow (P_cPP_d)(x)>0.$$

Thus, S is strictly positive. We can now apply the previous lemma to S to see that $u=te$ and $v=\frac{1}{t}e$ for some positive scalar t. From (3.1), these yield $b^2=td^2$ and $a^2=\frac{1}{t}c^2$. As all the quantities involved are positive, we see that b is a positive scalar multiple of d and a is a positive scalar multiple of c. Thus, a and b are unique up to positive scalar mutliples.

Now for the existence. We show that there exist $c,d>0$ such that P_cPP_d is doubly stochastic. Consider the compact convex set

$$K:=\{x\in\mathcal{V}_+:\operatorname{tr}(x)=1\}.$$

As the inverse map $x\mapsto x^{-1}$ is continuous on V_{++} and P is strictly positive, we see that the (nonlinear) maps $f:K\to\mathcal{V}_{++}$ and $F:K\to K$ defined by

$$f(x):=\Big[P^*\Big(P(x)^{-1}\Big)\Big]^{-1}\quad\text{and}\quad F(x):=\frac{f(x)}{\operatorname{tr}(f(x))}$$

are continuous. Hence, by the Brouwer fixed point theorem, there exists $v\in K$ such that $F(v)=v$, that is,

$$\Big[P^*\Big(P(v)^{-1}\Big)\Big]^{-1}=t\,v,\tag{3.2}$$

where $t=\operatorname{tr}(f(v))$. Define $u:=P(v)^{-1}$ so that $P(v)=u^{-1}=P_{\sqrt{u}^{-1}}(e)=\big(P_{\sqrt{u}}\big)^{-1}(e)$ and $P_{\sqrt{u}}P(v)=e$. As $v=P_{\sqrt{v}}(e)$, this becomes

$$P_{\sqrt{u}}PP_{\sqrt{v}}(e)=e.$$

Now, (3.2) implies $\big[P^*(u)\big]^{-1}=tv$, which simplifies to

$$P_{\sqrt{v}}P^*P_{\sqrt{u}}(e)=\frac{1}{t}e.$$

Writing $S:=P_{\sqrt{u}}PP_{\sqrt{v}}$, we see that $S(e)=e$ and $S^*(e)=\frac{1}{t}e$. As

$$n=\langle S(e),e\rangle=\langle e,S^*(e)\rangle=\frac{n}{t},$$

we get $t = 1$. Hence, $S(e) = e = S^*(e)$. Since S is a positive map, we conclude that S is doubly stochastic. Now we can put $c := \sqrt{u}$ and $d := \sqrt{v}$ to see that $c, d > 0$ and $P_c P P_d$ doubly stochastic. This completes the proof of Theorem 1.1. $\square$

4. Concluding remarks

We now relate Theorem 1.1 to a known result and formulate an open problem.
Let M_n denote the space of all $n \times n$ complex matrices. Suppose P is a positive linear map on M_n. This means that P is a (complex) linear map from M_n to itself, $P(\mathcal{H}^n) \subseteq \mathcal{H}^n$, and $P : \mathcal{H}^n \to \mathcal{H}^n$ is a (real linear) positive map. (We note that every positive map on $\mathcal{H}^n$ can be extended to a complex linear map on M_n in an obvious way, see e.g., Lemma 2.3.1 in [1].) With P positive on M_n, we say that P is said to be doubly stochastic if $P(I) = I = P^*(I)$, where I denotes the identity matrix in M_n.

The following result is due to Gurvitz [7]; see [9], Theorem 1.2, where it is called the *weak form of Gurvitz's generalization to positive maps.*

Theorem 4.1. *Let $P : M_n \to M_n$ be a linear map taking positive definite matrices to positive definite matrices. Then, there exist invertible matrices C and D such that the map $X \mapsto CP(DXD^*)C^*$ is doubly stochastic.*

Now, on $\mathcal{H}^n$, a map of the form $T : X \mapsto CXC^*$ with C invertible is a cone-automorphism, that is, $T(\mathcal{H}^n_+) = \mathcal{H}^n_+$. (Note: The map $X \mapsto C\overline{X}C^*$ is also a cone-automorphism, see [11].) On a Euclidean Jordan algebra $\mathcal{V}$, a linear map T is said to be a cone-automorphism if $T(\mathcal{V}_+) = \mathcal{V}_+$. Motivated by the above result of Gurvtiz, we formulate the following

Problem: On a Euclidean Jordan algebra $\mathcal{V}$, consider a linear map $P : \mathcal{V} \to \mathcal{V}$ that takes $\mathcal{V}_{++}$ to itself. Are there cone-automorphisms T and S on $\mathcal{V}$ such that TPS is doubly stochastic?

References

[1] R. Bhatia, Matrix Analysis, Springer-Verlag, New York, 1997.
[2] J. Faraut and A. Korányi, Analysis on Symmetric Cones, Oxford University Press, Oxford, 1994.
[3] T.T. Georgiou, M. Pavon, Positive contraction mappings for classical and quantum Schrödinger system, J. Mathematical Phy. 56 (2015) 033301, arXiv:1405.6650v3.
[4] M.S. Gowda, Positive and doubly stochastic maps, and majorization in Euclidean Jordan algebras, Linear Algebra Appl. 528 (2017) 40-61.
[5] M.S. Gowda, R. Sznajder, J. Tao, Some P-properties for linear transformations on Euclidean Jordan algebras, Linear Algebra Appl. 393 (2004) 203-232.
[6] M.S. Gowda, J. Tao, The Cauchy interlacing theorem in simple Euclidean Jordan algebras and some consequences, Linear Multi. Algebra, 59 (2011) 65-86.
[7] L. Gurvitz, Classical deterministic complexity of Edmonds' problem and quantum entanglement, In: Proceedings of the Thirty-fifth Annual ACM Symposium on Theory of Computing, (2003) pp. 10–19, arXiv:quant-ph/0303055.
[8] M. Idel, On the structure of positive maps, MA thesis, Ludwig-Maximilian Universität München, Technische Universität München, 2013.
[9] M. Idel, A review of matrix scaling and Sinkhorn's normal form for matrices and positive maps, arxiv: 1609.06349v1, 2016.

[10] M.V. Menon, Reduction of a matrix with positive elements to a doubly stochastic matrix, Proc. Amer. Math. Soc. 18 (1967) 244-247.
[11] H. Schneider, Positive operators and an inertia theorem, Numer. Math. 7 (1965) 11–17.
[12] R. Sinkhorn, A Relationship between arbitrary positive matrices and doubly stochastic matrices, Ann. Math. Stat. 35.2 (1964) 876-879.