Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Expo. Math. 31 (2013) 295–303
EXPOSITIONES MATHEMATICAE
www.elsevier.com/locate/exmath

Positive projections of symmetric matrices and Jordan algebras

Bent Fuglede*, Søren Tolver Jensen[1]

Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen Ø, Denmark

Received 1 August 2012; received in revised form 20 August 2012

Résumé

An elementary proof is given that the projection from the space of all symmetric $p \times p$ matrices onto a linear subspace is positive if and only if the subspace is a Jordan algebra. This solves a problem in a statistical model.

MSC 2010: primary 15A03; secondary 15A63

Keywords: Hyperorthogonal

1. Introduction

The result of the paper gives a solution to a problem in a statistical model. In [7, Theorem 3.1], it was proved that the orthogonal projection of the set $\mathcal{S}_p$ of symmetric $p \times p$-matrices onto a Jordan subalgebra $\mathcal{M}$ is a positive projection (the 'if part' of Theorem 1), and a conjecture of the opposite statement was presented (the 'only if part' of Theorem 1). The concept of a Jordan algebra goes back to Jordan [8]; see also [1,5].

It turns out that the validity of the conjecture follows from a more general result in [2, Theorem 1.4]. And in fact, the conjecture follows directly from the inequality of Kadison [9] in the finite dimensional case; see Remark 1. For further information, see [4,10].

* Corresponding author.
E-mail address: fuglede@math.ku.dk (B. Fuglede).
[1] Søren Tolver Jensen passed away in March 2012 after a lengthy illness. By that time the present manuscript was essentially complete.

doi:10.1016/j.exmath.2013.01.005

Although the theorem seems to be well known (Størmer, personal communication), we have not been able to find the theorem explicitly stated in the literature. Concerning the 'if part' it is proved in [2, Lemma 2.3] that the orthogonal projection is positive if $\mathcal{M}$ is a spin factor, and our proof of the 'if part' in the general finite dimensional case is essentially the same.

We shall present an elementary proof also of the 'only if part' of Theorem 1. A major step is the reduction to the commutative case. Our proof is furthermore based on the concept of an *hyperorthogonal* p-tuple of vectors in $\mathbb{R}^n$ (Definition 1, see also Theorem 2 in the closing section of the paper).

We thank Professor Erling Størmer for valuable comments to the manuscript.

2. The main result

For given $p \in \mathbb{N}$ we denote by $\mathcal{S}_p$ the linear space of all symmetric $p \times p$ matrices with real entries. The dimension $\dim \mathcal{S}_p = N$ equals $\frac{1}{2}p(p+1)$. An inner product on $\mathcal{S}_p$ is defined by

$$\langle A, B\rangle = \mathrm{Tr}(AB) = \mathrm{Tr}(BA), \quad A, B \in \mathcal{S}_p.$$

Consider a (linear) subspace $\mathcal{M}$ of $\mathcal{S}_p$ containing the unit matrix $I = (\delta_{ij})$, and write $\dim \mathcal{M} = m$ ($\leq N$). By $P_{\mathcal{M}}$ we denote the operator of (orthogonal) projection from $\mathcal{S}_p$ onto $\mathcal{M}$. The positive cone $\mathcal{S}_p^+$ consists of all positive semidefinite matrices in $\mathcal{S}_p$. We say that $P_{\mathcal{M}}$ is *positive* and write $P_{\mathcal{M}} \geq 0$ if $P_{\mathcal{M}}(\mathcal{S}_p^+) \subseteq \mathcal{S}_p^+$, that is if

$$P_{\mathcal{M}}(S) \in \mathcal{S}_p^+ \quad \text{for every } S \in \mathcal{S}_p^+. \tag{1}$$

Theorem 1. *$P_{\mathcal{M}} \geq 0$ holds if and only if $\mathcal{M}$ is a Jordan algebra, that is, $A^2 \in \mathcal{M}$ for every $A \in \mathcal{M}$.*

This is obvious for $m = 1$ because, on the one hand, $I^2 = I \in \mathcal{M}$, so $\mathcal{M}$ is Jordan, and on the other hand, $P_{\mathcal{M}}(S) = p^{-1}\mathrm{Tr}(S)I \in \mathcal{S}_p^+$ if $S \in \mathcal{S}_p^+$. Henceforth we assume that $m \geq 2$ and hence $p \geq 2$.

For any matrix $S = (s_{ij}) \in \mathcal{S}_p$ we denote by $S(x), x \in \mathbb{R}^p$, the quadratic form associated with S:

$$S(x) = \sum_{i,j=1}^{p} s_{ij}x_i x_j, \quad x = (x_1, \ldots, x_p) \in \mathbb{R}^p.$$

The condition $P_{\mathcal{M}} \geq 0$ is then characterized as follows.

Lemma 1. *Let $(A_1, \ldots, A_m)$ be an orthonormal base of $\mathcal{M}$. Then $P_{\mathcal{M}} \geq 0$ holds if and only if*

$$\sum_{k=1}^{m} A_k(x)A_k(y) \geq 0 \quad \textit{for } x, y \in \mathbb{R}^p. \tag{2}$$

Proof. For $S \in \mathcal{S}_p^+$ write

$$S = \sum_{i=1}^{p} \lambda_i E_i,$$

where $\lambda_i \geq 0$ are the eigenvalues of S, and E_i are pairwise orthogonal 1-dimensional projections. This spectral decomposition of S shows that it suffices to verify (1) for all 1-dimensional projections S. Any 1-dimensional projection has the form

$$E = (x_i x_j)_{i,j=1,\dots,p},$$

where $x = (x_1, \dots, x_p) \in \mathbb{R}^p$ has norm $|x| = \sqrt{x_1^2 + \cdots + x_p^2} = 1$. Note that $P_{\mathcal{M}}(E) = \sum_{k=1}^{m} \operatorname{Tr}(A_k E) A_k$. For any $A = (a_{ij}) \in \mathcal{S}_p$ we have

$$\operatorname{Tr}(AE) = \sum_{k=1}^{p} a_{ij} x_i x_j = A(x).$$

It follows for $y \in \mathbb{R}^p$ that

$$P_{\mathcal{M}}(E)(y) = \sum_{k=1}^{m} \operatorname{Tr}(A_k E) A_k(y) = \sum_{k=1}^{m} A_k(x) A_k(y), \tag{3}$$

which indeed is ≥ 0 for all rank 1 symmetric matrices $E = (x_i x_j)$ ($x \in \mathbb{R}^p$) and for all $y \in \mathbb{R}^p$, if and only if (2) holds. □

3. An elementary result

Proof of the 'if part' of Theorem 1. Supposing that $\mathcal{M}$ is a Jordan algebra (and that $m, p \geq 2$) we shall prove that $P_{\mathcal{M}} \geq 0$. For given $S \in \mathcal{S}_p^+$ the projection $P_{\mathcal{M}}(S) \in \mathcal{M} \subseteq \mathcal{S}_p$ is determined by

$$\operatorname{Tr}\big((S - P_{\mathcal{M}}(S))B\big) = 0 \quad \text{for all } B \in \mathcal{M}.$$

When $\mathcal{M}$ is Jordan and $A \in \mathcal{M}$ we may take $B = A^2$ to obtain

$$\operatorname{Tr}(A P_{\mathcal{M}}(S) A) = \operatorname{Tr}(P_{\mathcal{M}}(S) A^2) = \operatorname{Tr}(S A^2) = \operatorname{Tr}(ASA). \tag{4}$$

This leads to $P_{\mathcal{M}}(S) \in \mathcal{S}_p^+$ as follows.

Let $P_{\mathcal{M}}(S) = U^{-1} \Lambda U$ where $U \in O(p)$ is an orthogonal matrix and $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_p)$ is the diagonal matrix of eigenvalues of $P_{\mathcal{M}}(S)$, $\lambda_1 = \cdots = \lambda_d$ being the smallest eigenvalue of $P_{\mathcal{M}} S$ and of multiplicity d. If λ_1 were the only eigenvalue of $P_{\mathcal{M}}(S)$, that is if $d = p$, then $\lambda_1 \geq 0$, for if $\lambda_1 < 0$ then $P_{\mathcal{M}}(S) = \lambda_1 I$ and hence $\operatorname{Tr}(A P_{\mathcal{M}}(S) A) = \lambda_1 Tr(A^2) < 0$ for non-zero $A \in \mathcal{M}$ (for example $A = I$), in contradiction with (4) because $\operatorname{Tr}(ASA) \geq 0$ (since $S \in \mathcal{S}_p^+$). Thus actually $d < p$.

Let L be the Lagrange interpolation polynomial over $\mathbb{R}$ given by

$$L(X) = \frac{(X - \lambda_{d+1}) \cdots (X - \lambda_p)}{(\lambda_1 - \lambda_{d+1}) \cdots (\lambda_1 - \lambda_p)}. \tag{5}$$

Then $L(\lambda_1) = 1$ and $L(\lambda_j) = 0, j = d+1, \ldots, p$. Because $\mathcal{M}$ is Jordan and contains I we have $U^{-1}L(\Lambda)U = L(U^{-1}\Lambda U) = L(P_{\mathcal{M}}(S)) \in \mathcal{M}$, and by (4) applied to $A = U^{-1}L(\Lambda)U$

$$0 \le \mathrm{Tr}((U^{-1}L(\Lambda)U)S(U^{-1}L(\Lambda)U)) = \mathrm{Tr}((U^{-1}L(\Lambda)U)U^{-1}\Lambda U(U^{-1}L(\Lambda)U))$$
$$= \mathrm{Tr}(U^{-1}L(\Lambda)\Lambda L(\Lambda)U) = \mathrm{Tr}(L(\Lambda)\Lambda L(\Lambda)) = \sum_{i=1}^{p}(L(\lambda_i))^2\lambda_i = d\lambda_1.$$

Since the smallest eigenvalue λ_1 of $P_{\mathcal{M}}(S) = U^{-1}\Lambda U$ is non-negative, it follows that $P_{\mathcal{M}}(S) \in \mathcal{S}_p^+$. □

For the proof of the ‘only if part’ of Theorem 1 we begin by reducing it to the case where $\mathcal{M}$ is *commutative*, by application of the following expression for the projection on the intersection of two (not necessarily orthogonal) subspaces of $\mathbb{R}^p$.

Lemma 2. *Let $\mathcal{P}$ and $\mathcal{Q}$ be two subspaces of a finite dimensional Hilbert space $\mathcal{H}$, and let P and Q denote the operators of orthogonal projection from $\mathcal{H}$ onto $\mathcal{P}$ and $\mathcal{Q}$, respectively. The operator R of orthogonal projection on $\mathcal{R} := \mathcal{P} \cap \mathcal{Q}$ is then given by*

$$Rx = \lim_{n\to\infty}(PQ)^n x = \lim_{n\to\infty}(QP)^n x, \quad x \in \mathcal{H}.$$

Proof. Write $x = y + z$ with $y \in \mathcal{R}$ and $z \in \mathcal{R}^\perp$. The restrictions P' and Q' of P and Q, respectively, to $\mathcal{R}^\perp$ are the orthogonal projections from $\mathcal{R}^\perp$ onto $\mathcal{P} \ominus \mathcal{R}$ and $\mathcal{Q} \ominus \mathcal{R}$, respectively, and these two subspaces of $\mathcal{R}^\perp$ have only 0 in common. In terms of the operator norm $\|\cdot\|$ it follows that $\|P'Q'\| < 1$, for if $\|P'Q'\| = 1$ there would exist $z \in \mathcal{R}^\perp$ with $z \neq 0$ and $|PQz| = |P'Q'z| = |z|$. It would then follow that $|z| = |PQz| \le |Qz| \le |z|$ hence that $|PQz| = |Qz|$, and so $Qz \in \mathcal{P}$ and of course $Qz \in \mathcal{Q}$. Thus $Qz \in \mathcal{R}$ and also $Qz = Q'z \in \mathcal{R}^\perp$, so $Qz = 0$, in contradiction with $|PQz| = |z| \neq 0$. This shows that indeed $\|P'Q'\| < 1$, and since $y \in \mathcal{R} = \mathcal{P} \cap \mathcal{Q}$ we obtain $PQy = Py = y$ and hence

$$(PQ)^n x = (PQ)^n y + (PQ)^n z = y + (P'Q')^n z \to y = Rx \quad \text{as } n \to \infty. \quad \square$$

Proof of the ‘only if part’ of Theorem 1. Supposing that $P_{\mathcal{M}} \ge 0$ (and $\dim \mathcal{M} = m \ge 2$) we shall prove that the (linear) subspace $\mathcal{M}$ of $\mathcal{S}_p$ is a Jordan algebra. Denote by $\mathcal{D}_p$ the subspace of $\mathcal{S}_p$ consisting of all diagonal $p \times p$ matrices, and by $\mathcal{D}_p^+ = \mathcal{D}_p \cap \mathcal{S}_p^+$ the positive cone in $\mathcal{D}_p$. For any subspace $\mathcal{N}$ of $\mathcal{S}_p$ let $\mathcal{N}_d = \mathcal{N} \cap \mathcal{D}_p$ denote the subspace of $\mathcal{N}$ consisting of all diagonal matrices in $\mathcal{N}$. As before, $P_{\mathcal{M}}$ denotes the projection from $\mathcal{S}_p$ onto $\mathcal{M}$, and we write $P_{\mathcal{M}} \ge 0$ if $P_{\mathcal{M}}(\mathcal{S}_p^+) \subseteq \mathcal{S}_p^+$.

In particular, for any $S \in \mathcal{S}_p$, $P_{\mathcal{D}_p}(S)$ is the diagonal matrix S_d formed by the diagonal entries of S. Clearly, $P_{\mathcal{D}_p} \ge 0$. Furthermore, $\mathcal{M}_d$ is Jordan if $\mathcal{M}$ is so. It is therefore to be expected that

$$P_{\mathcal{M}} \ge 0 \quad \text{implies } P_{\mathcal{M}_d} \ge 0. \tag{6}$$

And this is indeed the case because $\mathcal{M}_d = \mathcal{M} \cap \mathcal{D}_p$ and hence by Lemma 2 applied to $\mathcal{H} = \mathcal{S}_p$

$$P_{\mathcal{M}_d}(S) = \lim_{n\to\infty}(P_{\mathcal{M}}P_{\mathcal{D}_p})^n(S) \in \mathcal{S}_p^+ \quad \text{for any } S \in \mathcal{S}_p^+,$$

$\mathcal{S}_p^+$ being closed in $\mathcal{S}_p$.

Now suppose that the remaining 'only if part' of Theorem 1 has been established for the particular case that $\mathcal{M}$ consists solely of diagonal matrices (this is essentially equivalent to $\mathcal{M}$ being commutative). For an arbitrary subspace $\mathcal{M}$ of $\mathcal{S}_p$ let $A \in \mathcal{M}$ be given, and let us prove that $A^2 \in \mathcal{M}$. There is an orthogonal matrix $U \in O(p)$ such that $U^{-1}AU$ is diagonal. Then $U^{-1}\mathcal{M}U$ is a subspace of $U^{-1}\mathcal{S}_pU = \mathcal{S}_p$ containing I. For any $S \in \mathcal{S}_p^+$ the projection $S' = P_{U^{-1}\mathcal{M}U}(S)$ satisfies $US'U^{-1} = P_{\mathcal{M}}(USU^{-1}) \in \mathcal{S}_p^+$ because $P_{\mathcal{M}} \geq 0$ and $USU^{-1} \in \mathcal{S}_p^+$. Thus $S' \in \mathcal{S}_p^+$, and so $P_{U^{-1}\mathcal{M}U} \geq 0$. It follows by (6) applied with $\mathcal{M}$ replaced by $U^{-1}\mathcal{M}U$ that $P_{(U^{-1}\mathcal{M}U)_d} \geq 0$. By hypothesis, $(U^{-1}\mathcal{M}U)_d$ is therefore Jordan, and consequently the diagonal matrix $U^{-1}AU \in (U^{-1}\mathcal{M}U)_d$ has the square $(U^{-1}AU)^2 = U^{-1}A^2U \in (U^{-1}\mathcal{M}U)_d \subseteq U^{-1}\mathcal{M}U$, that is $A^2 \in \mathcal{M}$, as claimed.

For the completion of the proof of the 'only if part' of Theorem 1 it thus remains to show that if $\mathcal{M}$ denotes a subspace of $\mathcal{S}_p$ consisting entirely of *diagonal matrices* (that is, if $\mathcal{M} \subset \mathcal{D}_p$) and if $P_{\mathcal{M}} \geq 0$ then $\mathcal{M}$ is Jordan. The restriction of $P_{\mathcal{M}}$ from $\mathcal{S}_p$ to $\mathcal{D}_p$ is of course likewise positive. Furthermore, $\mathcal{D}_p \cong \mathbb{R}^p$ when we identify a diagonal matrix $\operatorname{diag}(s_1, \ldots, s_p) \in \mathcal{D}_p$ with the vector $(s_1, \ldots, s_p) \in \mathbb{R}^p$. From now on we therefore *change the previous notation* by replacing $\mathcal{M} \subseteq \mathcal{D}_p$ with $\mathcal{M} \subseteq \mathbb{R}^p$ and by replacing $(\mathcal{D}_p)_+$ with $\mathbb{R}_+^p$, the positive cone in $\mathbb{R}^p$. Denote by $\mathcal{M}^\perp$ the orthogonal complement of $\mathcal{M}$ in $\mathbb{R}^p$. As before, $\dim \mathcal{M} = m \geq 2$, and we now write $\dim \mathcal{M}^\perp = n$, whereby $m + n = p$. Theorem 1 is trivial for $\mathcal{M} = \mathbb{R}^p$, and we may therefore assume that $n \geq 1$ and hence $p \geq 3$.

Consider an $n \times p$ matrix whose rows $b_l = (b_{l1}, \ldots, b_{lp})$ are linearly independent vectors in $\mathcal{M}^\perp$. The corresponding columns are denoted by $v_i = (v_{i1}, \ldots, v_{in})$, where $v_{il} = b_{li}$. For any vector $s = (s_1, \ldots, s_p) \in \mathcal{M}$ we have

$$\langle s, b_l\rangle = \sum_{i=1}^p s_i b_{li} = \sum_{i=1}^p s_i v_{il} = 0 \quad \text{for } l \in \{1, \ldots, n\},$$

that is,

$$\sum_{i=1}^p s_i v_i = 0 \quad \text{for every } s = (s_1, \ldots, s_p) \in \mathcal{M}. \tag{7}$$

Now suppose that $(b_1, \ldots, b_n)$ is even an *orthonormal base* of $\mathcal{M}^\perp$. Extend the normalized identity vector

$$a_1 = \frac{1}{\sqrt{p}}(1, \ldots, 1) \in \mathcal{M}$$

to an orthonormal base $(a_1, \ldots, a_m)$ for $\mathcal{M}$, and write

$$a_k = (a_{k1}, \ldots, a_{kp}) \in \mathbb{R}^p, \quad k \in \{1, \ldots, m\}.$$

We then have the orthonormal base $(a_1, \ldots, a_m, b_1, \ldots, b_n)$ for $\mathcal{M} \oplus \mathcal{M}^\perp = \mathbb{R}^p$. Consider the orthogonal $p \times p$ matrix Ω with rows $a_1, \ldots, a_m, b_1, \ldots, b_n$. For $i \in \{1, \ldots, p\}$ write

$$u_i = (a_{1i}, \ldots, a_{mi}) \in \mathbb{R}^m, \qquad v_i = (b_{1i}, \ldots, b_{ni}) \in \mathbb{R}^n. \tag{8}$$

The i'th column of Ω is formed by u_i followed by v_i, $i = 1, \ldots, p$. Then

$$\langle u_i, u_j\rangle + \langle v_i, v_j\rangle = \delta_{ij}, \quad i, j \in \{1, \ldots, p\},$$

and hence by Lemma 1 applied to $A_k = \mathrm{diag}(a_{k1}, \ldots, a_{kp})$ and to $x = e_i$, $y = e_j$ (where $e_1, \ldots, e_p$ denote the standard basic vectors in $\mathbb{R}^p$), noting that $A_k(e_i) = a_{ki}$:

$$-\langle v_i, v_j\rangle = \langle u_i, u_j\rangle = \sum_{k=1}^{m} a_{ki}a_{kj} = \sum_{k=1}^{m} A_k(e_i)A_k(e_j) \geq 0 \quad \text{for } i \neq j. \tag{9}$$

Thus the p-tuple $(v_1, \ldots, v_p)$ in $\mathbb{R}^n$ is *hyperorthogonal* in the following sense.

Definition 1. Let $p, n \in \mathbb{N}$. A p-tuple $(v_1, \ldots, v_p)$ of vectors in $\mathbb{R}^n$ is said to be *hyperorthogonal* if $\langle v_i, v_j\rangle \leq 0$ for any distinct $i, j \in \{1, \ldots, p\}$, that is, if the angle between any two distinct non-zero v_i and v_j (if there are such) is no less than $\pi/2$.

Thus the vectors v_i are neither required to be distinct nor to be non-zero. However, the non-zero vectors v_i must clearly be distinct, and at most two of them can be real multiples of the same vector. Any orthogonal p-tuple is of course hyperorthogonal. The $2n$-tuples of non-zero hyperorthogonal vectors in $\mathbb{R}^n$ are explicitly described in Theorem 2 at the end of the paper (but that description is not used in the present proof).

Lemma 3. *For any subspace $\mathcal{M}$ of $\mathbb{R}^p$, the projection from $\mathbb{R}^p$ onto $\mathcal{M}$ is positive if and only if the p-tuple $(v_1, \ldots, v_p)$ from (8) is hyperorthogonal.*

Proof. The 'only if part' of this lemma was established in (9) as a consequence of Lemma 1. Conversely, suppose $(v_1, \ldots, v_p)$ is hyperorthogonal, that is by (9), $\sum_{k=1}^{m} a_{ki}a_{kj} \geq 0$ for $i, j \in \{1, \ldots, p\}$, the case $i = j$ being trivial. It follows that

$$\sum_{k=1}^{m} A_k(x)A_k(y) = \sum_{i,j=1}^{p} \left(\sum_{k=1}^{m} a_{ki}a_{kj}\right) x_i^2 y_j^2 \geq 0$$

for $x, y \in \mathbb{R}^p$, and this implies $P_{\mathcal{M}} \geq 0$, again by Lemma 1. □

Returning to the proof of the 'only if part' of Theorem 1, the hypothesis is that $(v_1, \ldots, v_p)$ is hyperorthogonal. Not all v_i can be 0 because the corresponding rows b_l are normalized and in particular non-zero. We may therefore assume for example that $v_p \neq 0$. By projecting $v_1, \ldots, v_p$ onto $(\mathbb{R}v_p)^\perp$ within $\mathbb{R}^n$ we obtain another p-tuple $(v'_1, \ldots, v'_p)$ in $\mathbb{R}^n$, whereby $v'_p = 0$. Put $w = v_p/|v_p|$. Then $v'_i = v_i - \langle v_i, w\rangle w$ and $v'_j = v_j - \langle v_j, w\rangle w$, and for distinct indices $i, j \in \{1, \ldots, p-1\}$ we have $\langle v'_i, v'_j\rangle = \langle v_i, v_j\rangle - \langle v_i, w\rangle\langle v_j, w\rangle \leq 0$, because $\langle v_i, v_j\rangle \leq 0$, $\langle v_i, w\rangle \leq 0$, and $\langle v_j, w\rangle \leq 0$. Hence the $(p-1)$-tuple $(v'_1, \ldots, v'_{p-1})$ is hyperorthogonal in $\mathbb{R}^n$, and so is therefore the p-tuple $(v'_1, \ldots, v'_p)$ because $v'_p = 0$. From (7) we get by projection on $(\mathbb{R}v_p)^\perp$

$$\sum_{i=1}^{p} s_i v'_i = 0 \quad \text{for every } s = (s_1, \ldots, s_p) \in \mathcal{M}. \tag{10}$$

If $(v'_1, \ldots, v'_{p-1}) \neq (0, \ldots, 0)$ in $\mathbb{R}^n$, say $v'_{p-1} \neq 0$, we project the hyperorthogonal p-tuple $(v'_1, \ldots, v'_p)$ (where $v'_p = 0$) onto $(\mathbb{R}v'_{p-1})^\perp$ within $\mathbb{R}^n$ and obtain a hyperorthogonal

p-tuple $(v''_1, \dots, v''_p)$ in $\mathbb{R}^n$. Note that $v''_{p-1} = v''_p = 0$. This process of repeated projection ends after at most $p-1$ steps with a hyperorthogonal p-tuple $(z_1, \dots, z_p) \neq (0, \dots, 0)$ in $\mathbb{R}^n$, say with $z_j \neq 0$, such that the projections of $z_1, \dots, z_p$ on $(\mathbb{R}z_j)^\perp$ are all zero. Hence there are numbers $c_i \in \mathbb{R}$ such that $z_i = c_i z_j$ for $i \neq j$. As in (10) we obtain

$$\sum_{i=1}^{p} s_i z_i = 0 \quad \text{for every } s = (s_1, \dots, s_p) \in \mathcal{M}. \tag{11}$$

Since the p-tuple $z_1, \dots, z_p$ is hyperorthogonal, the inequality $c_i \neq 0$ holds for at most one index $i \neq j$. Actually, $c_i \neq 0$ holds for precisely one index $i \neq j$ in view of (11) applied to $s = (1, \dots, 1)$. For simplicity of writing we assume that this index i is 1 and that $j = 2$. Then $z_3 = 0, \dots, z_p = 0$ and (11) reads $s_1 z_1 + s_2 z_2 = 0$ for every $s \in \mathcal{M}$. For $s = (1, \dots, 1)$ this becomes $z_1 + z_2 = 0$. Since $(z_1, z_2) \neq (0, 0)$ it therefore follows from (11) that

$$s_1 = s_2 \quad \text{for every } s = (s_1, \dots, s_p) \in \mathcal{M}. \tag{12}$$

Defining the injective linear map $T = (T_1, \dots, T_p) : \mathbb{R}^{p-1} \to \mathbb{R}^p$ by

$$T(x_1, x_2, \dots, x_{p-1}) = (x_1, x_1, x_2, \dots, x_{p-1})$$

we thus have $\mathcal{M} \subseteq T(\mathbb{R}^{p-1})$, and the pre-image $\mathcal{N} = T^{-1}(\mathcal{M})$ is a subspace of $\mathbb{R}^{p-1}$ and is mapped bijectively by T onto $\mathcal{M}$. We show that $P_\mathcal{M} \geq 0$ implies $P_\mathcal{N} \geq 0$ (this is to be expected because $\mathcal{N}$ obviously is Jordan if $\mathcal{M}$ is so). Since $\dim \mathcal{N} = \dim \mathcal{M} = m$ we have $\dim \mathcal{N}^\perp = p - 1 - m = n - 1$. Let $b_1^*, \dots, b_{n-1}^*$ be an orthonormal base of $\mathcal{N}^\perp$ $(\subset \mathbb{R}^{p-1})$, and write in coordinates $b_l^* = (b_{l1}^*, \dots, b_{l,p-1}^*)$. Define new vectors $b_l = (0, b_{l1}^*, \dots, b_{l,p-1}^*)$ for $l \in \{1, \dots, n-1\}$. Clearly, $\langle b_l, T(s) \rangle = \langle b_l^*, s \rangle = 0$ for any $s = (s_1, \dots, s_{p-1}) \in \mathcal{N}$, that is, for $T(s) \in \mathcal{M}$, and so $b_l \in \mathcal{M}^\perp$ for $l \in \{1, \dots, n-1\}$.

Like the b_l^* in $\mathcal{N}^\perp$, the b_l form an orthonormal $(n-1)$-tuple in $\mathcal{M}^\perp$, and since $\dim \mathcal{M}^\perp = n$ an orthonormal base of $\mathcal{M}^\perp$ is obtained from $(b_1, \dots, b_{n-1})$ by adjoining a single normalized vector $b_n = (b_{n1}, \dots, b_{np}) \in \mathcal{M}^\perp$ orthogonal to $b_1, \dots, b_{n-1}$. By Lemma 3 the new (column) vectors $v_i = (b_{1i}, \dots, b_{ni})$, $i \in \{1, \dots, p\}$, cf. (8), satisfy (9), that is

$$\langle v_i, v_j \rangle = b_{1i} b_{1j} + \cdots + b_{ni} b_{nj} \leq 0 \quad \text{for distinct } i, j \in \{1, \dots, p\}. \tag{13}$$

For $j = 1$ and $i > 1$ this implies that $b_{ni} b_{n1} \leq 0$ since $b_{11}, \dots, b_{n-1,1} = 0$.

Suppose first that $b_{n1} = 0$ and hence $b_{l1} = 0$ for all $l \in \{1, \dots, n\}$, which means that the first basic vector $e_1 = (1, 0, \dots, 0)$ in $\mathbb{R}^p$ is orthogonal to $b_1, \dots, b_n$ and thus belongs to $\mathcal{M}^{\perp\perp} = \mathcal{M}$ in contradiction with (12).

Hence $b_{n1} \neq 0$ and it follows from (13) with $j = 1$ that b_{ni} for $i \in \{2, \dots, p\}$ all have the same sign (if unequal to 0), and hence $b_{ni} b_{nj} \geq 0$ for (distinct) $i, j > 1$. Corresponding to (8) (now with p replaced by $p-1$) we define

$$v_i^* = (b_{1i}^*, \dots, b_{n-1,i}^*) = (b_{1,i+1}, \dots, b_{n-1,i+1}) \in \mathbb{R}^{n-1}$$

for $i \in \{1, \dots, p-1\}$. Consequently,

$$\langle v_i^*, v_j^* \rangle = \langle v_{i+1}, v_{j+1} \rangle - b_{n,i+1} b_{n,j+1} \leq 0$$

for distinct $i, j \in \{1, \dots, p-1\}$, and the $(p-1)$-tuple $(v_1^*, \dots, v_{p-1}^*)$ on $\mathbb{R}^{n-1}$ is therefore hyperorthogonal along with the new p-tuple $(v_1, \dots, v_p)$ in $\mathbb{R}^n$. According to Lemma 3 this shows that indeed $P_{\mathcal{N}} \geq 0$. By induction with respect to p we may again assume that the 'only if part' of Theorem 1 holds when p is replaced by $p-1$. Thus $\mathcal{N}$ is Jordan, and so is therefore $\mathcal{M} = T(\mathcal{N})$. □

Remarks.

Remark 1. As in [2], the proof of the 'only if part' can be obtained right away by the inequality of Kadison [9], which asserts that if the orthogonal projection $P_{\mathcal{M}}$ is positive, then

$$P_{\mathcal{M}}(A^2) \geq (P_{\mathcal{M}}(A))^2 \quad \text{for every } A \in \mathcal{S}_p. \tag{14}$$

For if $A \in \mathcal{M}$, then $P_{\mathcal{M}}(A) = A$ and $P_{\mathcal{M}}(A^2) \geq A^2$. Since

$$\mathrm{Tr}(P_{\mathcal{M}}(A^2)) = \langle P_{\mathcal{M}}(A^2), I \rangle = \langle A^2, I \rangle = \mathrm{Tr}(A^2)$$

it follows that $A^2 = P_{\mathcal{M}}(A^2) \in \mathcal{M}$.

As noticed by Kadison [9, p. 500], the inequality (14) is the ordinary Schwarz inequality in the commutative case.

Remark 2. Let $\mathcal{H}_p$ denote the $\mathbb{R}$-linear space of all hermitian $p \times p$ matrices with complex entries. In a natural way, $\mathcal{H}_p$ is a Jordan subalgebra of $\mathcal{S}_{2p}$. Hence the orthogonal projection of $\mathcal{S}_{2p}$ onto $\mathcal{H}_p$ is positive, and a complex version of Theorem 1 is obtained for a *real* subspace $\mathcal{M}$ of $\mathcal{H}_p$.

The representation of a Jordan algebra as symmetric $p \times p$-matrices with *real* entries is given in [6], and based on this representation an explicit expression of the orthogonal projection is obtained.

In the special case of a spin factor the same representation is given by Jacobson [5].

Hyperorthogonal p-tuples on the unit sphere in $\mathbb{R}^n$.

We propose to determine all hyperorthogonal p-tuples (see Definition 1) of non-zero vectors, or just as well of normalized vectors, in $\mathbb{R}^n$, in the particular case where $p = 2n$. Let Σ_n denote the unit sphere in $\mathbb{R}^n$ $(n \geq 1)$, and d the standard distance on Σ_n. Hyperorthogonality of a p-tuple $(v_1, \dots, v_p)$ on Σ_n then amounts to $d(v_i, v_j) \geq \pi/2$ for distinct $i, j \in \{1, \dots, p\}$. Every 1-tuple is of course hyperorthogonal, so we assume that $p \geq 2$.

A pair (v_1, v_2) of points of Σ_n is termed an *antipodal pair* if $v_1 = -v_2$, in other words if $d(v_1, v_2) = \pi$.

Theorem 2. (a) *Every hyperorthogonal $2n$-tuple on Σ_n consists of n mutually orthogonal antipodal pairs. In other words, every hyperorthogonal $2n$-tuple is obtained from an orthonormal base $(v_1, \dots, v_n)$ for $\mathbb{R}^n$ by adjoining the opposite base $(-v_1, \dots, -v_n)$.*
(b) *There exists a hyperorthogonal p-tuple on Σ_n if and only if $p \leq 2n$.*

Thus, for $n = 3$, the only hyperorthogonal 6-tuple on the 2-sphere Σ_3 consists of the vertices of a regular octahedron inscribed in the unit ball in $\mathbb{R}^3$. For an explicit determination of all hyperorthogonal p-tuples on Σ_n for arbitrary $p \leq 2n$, see [3].

Proof of Theorem 2. For $n = 1$, (a) is obvious, so assume that $n \geq 2$. For the inductive proof suppose that (a) holds for smaller values of n.

As shown in the paragraph containing (10) the orthogonal projection of $(v_1, \ldots, v_{2n-1})$ on $(\mathbb{R}v_{2n})^{\perp}$ is a hyperorthogonal $(2n-1)$-tuple $(v'_1, \ldots, v'_{2n-1})$. At most one of the vectors $v_1, \ldots, v_{2n-1}$ could be proportional to v_{2n}. Therefore, at most one of the vectors $v'_1, \ldots, v'_{2n-1}$ is 0. We may assume for example that $v'_i \neq 0$ for $i = 1, \ldots, 2n-2$.

Let $w_i = v'_i/\|v'_i\|, i = 1, \ldots, 2n-2$. Then $(w_1, \ldots, w_{2n-2})$ is a hyperorthogonal $2n-2$ tuple on the 'equator' $\Sigma^* \cong \Sigma_{n-1}$ corresponding to the 'pole' v_{2n}:

$$\Sigma^* = \{x \in \Sigma_n : d(x, v_{2n}) = \pi/2\}. \tag{15}$$

By induction, this hyperorthogonal $(2n-2)$-tuple consists of $n-1$ mutually orthogonal antipodal pairs, say $(w_1, w_n), \ldots, (w_{n-1}, w_{2n-2})$, where $(w_1, w_2, \ldots, w_{n-1})$ is an orthonormal base for $(\mathbb{R}v_{2n})^{\perp}$ and $-w_i = w_{i+n-1},\ i = 1, \ldots, n-1$.

Since $\langle v'_{2n-1}, v'_i \rangle \leq 0$ for $i = 1, \ldots, 2n-2$, it follows that $\langle v'_{2n-1}, \pm w_i \rangle \leq 0$ for $i = 1, \ldots, n-1$ and therefore that $v'_{2n-1} = 0$. Since $\langle v_{2n-1}, v_{2n} \rangle \leq 0$ it follows that $v_{2n-1} = -v_{2n}$, and $\langle v_i, \pm v_{2n} \rangle \leq 0$ for $i = 1, \ldots, 2n-2$. Thus $\langle v_i, v_{2n} \rangle = 0$ for $i = 1, \ldots, 2n-2$, and $(v_1, \ldots, v_{2n-2}) = (v'_1, \ldots, v'_{2n-2}) = (w_1, \ldots, w_{2n-2})$, and it follows that $(v_1, \ldots, v_{2n}) = (w_1, \ldots, w_{2n-2}, v_{2n-1}, v_{2n})$ consists of n mutually orthogonal antipodal pairs.

The 'if part' of (b) is of course a consequence of (a). For the 'only if part' of (b), suppose that $(v_1, \ldots, v_{2n+1})$ is an hyperorthogonal $(2n+1)$-tuple of normalized vectors in $\mathbb{R}^n$. According to (a) we may suppose that $(v_1, \ldots, v_{2n})$ is as described in (a) (the latter description). Then $\langle v_{2n+1}, \pm v_i \rangle \leq 0$ for $i \leq n$, and so v_{2n+1} is orthogonal to each v_i, in contradiction with $|v_{2n+1}| = 1$. This completes the proof of Theorem 2. □

References

[1] A.A. Albert, On Jordan algebras of linear transformations, Trans. Amer. Math. Soc. 59 (1946) 524–555.

[2] E.D. Effros, E. Størmer, Positive projections and Jordan structure in operator algebras, Math. Scand. 45 (1979) 127–138.

[3] B. Fuglede, Hyperorthogonal families of vectors, 2012, Manuscript.

[4] U. Haagerup, E. Størmer, Positive projections of von Neumann algebras on JW-algebras, Rep. Math. Phys. 36 (1995) 317–330.

[5] N. Jacobson, Jordan algebras of real symmetric matrices, Algebras Groups Geom. 4 (1987) 291–304.

[6] S.T. Jensen, Covariance hypotheses which are linear in both the covariance and the inverse covariance, Ann. Statist. 16 (1988) 302–322.

[7] S.T. Jensen, J. Madsen, Estimation of proportional covariances in the presence of certain linear restrictions, Ann. Statist. 32 (2004) 219–232.

[8] P. Jordan, Über eine Klasse nichtassoziativer hyperkomplexer Algebren, Nachr. Ges. Wiss. Göttingen Math. Phys. Kl. (1932) 569–575.

[9] R.V. Kadison, A generalized Schwarz inequality and algebraic invariants for operator algebras, Ann. of Math. 56 (1952) 494–503.

[10] E. Størmer, Decomposition of positive projections on C^*-algebras, Math. Ann. 247 (1980) 21–41.