PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 33, Fasc. 2 (2013), pp. 191–199

SOME DECOMPOSITIONS OF MATRIX VARIANCES

BY

ZOLTÁN LÉKA (BUDAPEST) AND DÉNES PETZ (BUDAPEST)

*This paper is dedicated to Rajendra Bhatia
on the occasion of his 60th birthday*

Abstract. When D is a density matrix and A_1, A_2 are self-adjoint operators, then the standard variance is a 2×2 matrix:

$$\mathrm{Var}_D(A_1, A_2)_{i,j} := \mathrm{Tr}\, DA_iA_j - (\mathrm{Tr}\, DA_i)(\mathrm{Tr}\, DA_j) \quad (1 \leqslant i, j \leqslant 2).$$

The main result in this work is that there are projections P_k such that $D = \sum_k \lambda_k P_k$ with $0 < \lambda_k$ and $\sum_k \lambda_k = 1$ and $\mathrm{Var}_D(A_1, A_2) = \sum_k \lambda_k \mathrm{Var}_{P_k}(A_1, A_2)$. In a previous paper only the $A_1 = A_2$ case was included and the relevance is motivated by the paper [8].

2000 AMS Mathematics Subject Classification: Primary: 62J10; Secondary: 62F30.

Key words and phrases: Density matrix, variance, covariance, decomposition, projections.

1. INTRODUCTION

The subject of the paper is matrix theory, see [1] and [3]. By a *density matrix* $D \in M_n(\mathbb{C})$ we mean $D \geqslant 0$ and $\mathrm{Tr}\, D = 1$. In quantum information theory the traditional variance is

$$\mathrm{Var}_D(A) = \mathrm{Tr}\, DA^2 - (\mathrm{Tr}\, DA)^2 \tag{1.1}$$

when D is a density matrix and $A \in M_n(\mathbb{C})$ is a self-adjoint operator (see [3], [5], and [6]). This is a simple example, but here A_1, A_2 are self-adjoint operators. Then the standard variance is a matrix:

$$\mathrm{Var}_D(A_1, A_2) = \begin{bmatrix} \mathrm{Tr}\, DA_1^2 - (\mathrm{Tr}\, DA_1)^2 & \mathrm{Tr}\, DA_1A_2 - (\mathrm{Tr}\, DA_1)(\mathrm{Tr}\, DA_2) \\ \mathrm{Tr}\, DA_2A_1 - (\mathrm{Tr}\, DA_2)(\mathrm{Tr}\, DA_1) & \mathrm{Tr}\, DA_2^2 - (\mathrm{Tr}\, DA_2)^2 \end{bmatrix}.$$

Assume that $0 \leqslant \lambda_1, \lambda_2$ and $\lambda_1 + \lambda_2 = 1$. By an elementary computation we obtain

$$\mathrm{Var}_{\lambda_1 D_1 + \lambda_2 D_2}(A_1, A_2) - \lambda_1 \mathrm{Var}_{D_1}(A_1, A_2) - \lambda_2 \mathrm{Var}_{D_2}(A_1, A_2) = \lambda_1 \lambda_2 \begin{bmatrix} a^2 & ab \\ ab & b^2 \end{bmatrix} \geqslant 0,$$

where

$$a = \mathrm{Tr}\,(D_1 - D_2)A_1 \quad \text{and} \quad b = \mathrm{Tr}\,(D_1 - D_2)A_2.$$

It follows that we have the concavity of the variance functional $D \mapsto \mathrm{Var}_D(A_1, A_2)$:

$$\mathrm{Var}_D(A_1, A_2) \geqslant \sum_i \lambda_i \mathrm{Var}_{D_i}(A_1, A_2) \quad \text{if } D = \sum_i \lambda_i D_i,$$

where $\lambda_i \geqslant 0$ and $\sum_i \lambda_i = 1$. Here the equality may be also true and this is the result in Theorem 3.1: D is a certain convex combination of projections P_i as $D = \sum_i p_i P_i$ and

$$\mathrm{Var}_D(A_1, A_2) = \sum_i p_i \mathrm{Var}_{P_i}(A_1, A_2).$$

We note that our results can be interpreted in the recent terminology of roof (see, e.g., [9]) that originates from quantum theory. It seems to be important to understand roofs because they admit convex decompositions of various quantum mechanical quantities; see, e.g., [6], [8], [10]. In this context, the introduced variance matrix is a concave roof of itself.

The particular case $A_1 = A_2$ was already obtained in [6]. It is easy to show that

$$\mathrm{Var}_D(A_1 + \lambda_1 I, A_2 + \lambda_2 I) = \mathrm{Var}_D(A_1, A_2) \quad (\lambda_1, \lambda_2 \in \mathbb{R}).$$

Therefore we can assume $\mathrm{Tr}\, DA_1 = \mathrm{Tr}\, DA_2 = 0$.

2. GENERAL COMPUTATIONS

We are interested in the projections $P_1, P_2, \dots, P_N$ when given a density D and self-adjoint matrices A_1, A_2 such that $D = \sum_i \lambda_i P_i$ ($\lambda_i \geqslant 0$, $\sum_i \lambda_i = 1$) and

$$\mathrm{Var}_D(A_1, A_2) = \sum_i \lambda_i \mathrm{Var}_{P_i}(A_1, A_2). \tag{2.1}$$

First we make an elementary computation when $\sum_i \lambda_i P_i = D$. (It is not assumed that the projections P_i are orthogonal.) The point is to find a 2×2 matrix:

$$\mathrm{Var}_D(A_1, A_2) - \sum_i \lambda_i \mathrm{Var}_{P_i}(A_1, A_2) =: \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix}.$$

We compute α, β, γ:

$$\begin{aligned}\alpha &= -\big(\sum_i \lambda_i \mathrm{Tr}\, P_i A_1\big)^2 + \sum_i \lambda_i (\mathrm{Tr}\, P_i A_1)^2 \\ &= \sum_i \lambda_i (\mathrm{Tr}\, P_i A_1)(\mathrm{Tr}\, P_i A_1 - \mathrm{Tr}\, D A_1), \\ \gamma &= \sum_i \lambda_i (\mathrm{Tr}\, P_i A_2)(\mathrm{Tr}\, P_i A_2 - \mathrm{Tr}\, D A_2),\end{aligned}$$

$$\begin{aligned}\beta &= -\big(\sum_i \lambda_i \mathrm{Tr}\, P_i A_1\big)\big(\sum_j \lambda_j \mathrm{Tr}\, P_j A_2\big) + \sum_i \lambda_i (\mathrm{Tr}\, P_i A_1)(\mathrm{Tr}\, P_i A_2) \\ &= \sum_i \lambda_i (\mathrm{Tr}\, P_i A_1)(\mathrm{Tr}\, P_i A_2 - \mathrm{Tr}\, D A_2).\end{aligned}$$

Given the density D and the self-adjoint matrices A_1 and A_2 we should find the following solution:

(a) $D = \sum_i \lambda_i P_i$, where P_i's are projections, $\lambda_i \geqslant 0$, and $\sum_i \lambda_i = 1$;

(b) $\sum_i \lambda_i (\mathrm{Tr}\, P_i A_1)^2 = (\mathrm{Tr}\, D A_1)^2$ ($\alpha = 0$);

(c) $\sum_i \lambda_i (\mathrm{Tr}\, P_i A_2)^2 = (\mathrm{Tr}\, D A_2)^2$ ($\gamma = 0$);

(d) $\sum_i \lambda_i (\mathrm{Tr}\, P_i A_1)(\mathrm{Tr}\, P_i A_2 - \mathrm{Tr}\, D A_2) = 0$ ($\beta = 0$).

Instead of (d) we can take

(d′) $\sum_i \lambda_i \big(\mathrm{Tr}\, P_i (A_1 + A_2)\big)^2 = \big(\mathrm{Tr}\, D(A_1 + A_2)\big)^2$.

This implies that the above conditions (b)–(d) have an equivalent form.

THEOREM 2.1. *The condition* (2.1) *has the equivalent form*

$$\sum_i \lambda_i \big(\mathrm{Tr}\, P_i(\alpha A_1 + \beta A_2)\big)^2 = \big(\mathrm{Tr}\, D(\alpha A_1 + \beta A_2)\big)^2 \tag{2.2}$$

for $\alpha, \beta \in \{0, 1\}$.

We shall give a solution for the formula (2.2). First we have the following result when D has rank 2.

LEMMA 2.1. *Let* $D \in M_n(\mathbb{C})$ *be a density matrix with* $\mathrm{rank} D = 2$ *and let* $A_1, A_2 \in M_n(\mathbb{C})$ *be self-adjoint matrices such that* $\mathrm{Tr}\, D A_1 = \mathrm{Tr}\, D A_2 = 0$. *There exist projections* $P_1, P_2 \in M_n(\mathbb{C})$ *and* $p \in (0, 1)$ *such that*

$$D = pP_1 + (1 - p)P_2$$

and

$$\mathrm{Var}_D(A_1, A_2) = p\mathrm{Var}_{P_1}(A_1, A_2) + (1 - p)\mathrm{Var}_{P_2}(A_1, A_2).$$

Proof. Without loss of generality one can assume that D is diagonal, hence it is enough to prove the statement when $n = 2$. We recall that any 2×2 self-adjoint ρ with $\mathrm{Tr}\, \rho = 1$ can be written as the real linear combination of the identity

I and the Pauli matrices. In fact, let

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix};$$

then we have

$$\rho = \frac{1}{2}(I + x\sigma_x + y\sigma_y + z\sigma_z)$$

(this is called the *Bloch representation*); the self-adjoint points on the Bloch sphere, that is $x^2 + y^2 + z^2 = 1$, correspond to the pure states. Moreover, any traceless self-adjoint 2×2 matrix can be recovered by real linear combinations of the Pauli matrices as well.

Now one can write $(\mathrm{Tr}\, A_1)^{-1} A_1$ and $(\mathrm{Tr}\, A_2)^{-1} A_2$ in Bloch's representation (if $\mathrm{Tr}\, A_1 = 0$ or $\mathrm{Tr}\, A_2 = 0$, rewrite A_1 and A_2 as the linear combination of the Pauli matrices). Hence the densities which are orthogonal to A_1 and A_2 are lying at the non-empty intersection of two affine subspaces of $\mathbb{R}^3$. By the assumption $\mathrm{Tr}\, DA_1 = \mathrm{Tr}\, DA_2 = 0$, the intersection contains the point of the Bloch ball that represents D, and hence meets the Bloch sphere as well in P_1 and P_2. Then D is the convex combination of the projections P_1 and P_2. Since $\mathrm{Tr}\, A_1 P_i = \mathrm{Tr}\, A_2 P_i = 0$ holds, we readily obtain the decomposition of the variance as well. ■

3. DECOMPOSITION OF THE MATRIX VARIANCE

To prove the main theorem here, we need the following lemma.

LEMMA 3.1. *Let* $D \in M_n(\mathbb{C})$ *be a density matrix with* $\mathrm{rank} D \geqslant 3$ *and let* $A_1, A_2 \in M_n(\mathbb{C})$ *be self-adjoint matrices such that* $\mathrm{Tr}\, DA_1 = \mathrm{Tr}\, DA_2 = 0$. *There exist densities* $D_1, D_2 \in M_n(\mathbb{C})$ *and* $p \in (0,1)$ *such that*

$$D = pD_1 + (1-p)D_2, \qquad \mathrm{Tr}\, D_i A_1 = \mathrm{Tr}\, D_i A_2 \ = \ 0 \ (i = 1, 2),$$

and $\mathrm{rank} D_i < \mathrm{rank} D$ *for* $i = 1, 2$.

P r o o f. Let us assume the matrix form

$$D = \mathrm{Diag}(d_{-j}, \ldots, d_{-1}, d_0, d_1, \ldots, d_k),$$

where $j, k \geqslant 1$ and $d_{-1}, d_0, d_1 > 0$. Then we shall construct D_1 and D_2 in the block-matrix forms as

$$D_1 = \begin{bmatrix} \mathrm{Diag}(d_{-j}, \ldots, d_{-2}) & 0 & 0 & 0 & 0 \\ 0 & d_{-1} & x & y & 0 \\ 0 & x & d_0 & z & 0 \\ 0 & y & z & d_1 & 0 \\ 0 & 0 & 0 & 0 & \mathrm{Diag}(d_2, \ldots, d_k) \end{bmatrix}$$

and

$$D_2 = \begin{bmatrix} \mathrm{Diag}(d_{-j},\dots,d_{-2}) & 0 & 0 & 0 & 0 \\ 0 & d_{-1} & \frac{-px}{1-p} & \frac{-py}{1-p} & 0 \\ 0 & \frac{-px}{1-p} & d_0 & \frac{-pz}{1-p} & 0 \\ 0 & \frac{-py}{1-p} & \frac{-pz}{1-p} & d_1 & 0 \\ 0 & 0 & 0 & 0 & \mathrm{Diag}(d_2,\dots,d_k) \end{bmatrix},$$

where $x, y, z \in \mathbb{R}$. Let $D_1^>$ and $D_2^>$ denote the 3×3 matrices from D_1 and D_2:

$$D_1^> = \begin{bmatrix} d_{-1} & x & y \\ x & d_0 & z \\ y & z & d_1 \end{bmatrix}, \quad D_2^> = \begin{bmatrix} d_{-1} & \frac{-px}{1-p} & \frac{-py}{1-p} \\ \frac{-px}{1-p} & d_0 & \frac{-pz}{1-p} \\ \frac{-py}{1-p} & \frac{-pz}{1-p} & d_1 \end{bmatrix}.$$

Now $p \in (0,1)$ is a fixed parameter, its value will be determined later. First, we need to guarantee that

$$\mathrm{Tr}\, D_i A_1 = \mathrm{Tr}\, D_i A_2 = 0 \quad (i = 1, 2)\,.$$

Since $\mathrm{Tr}\, DA_1 = \mathrm{Tr}\, DA_2 = 0$, the last equalities can be written as

$$\mathrm{Tr}\, D_i A_1 = \mathrm{Re}\,\big(x(A_1)_{j,j+1} + y(A_1)_{j,j+2} + z(A_1)_{j+1,j+2}\big) = 0,$$
$$\mathrm{Tr}\, D_i A_2 = \mathrm{Re}\,\big(x(A_2)_{j,j+1} + y(A_2)_{j,j+2} + z(A_2)_{j+1,j+2}\big) = 0.$$

The (x,y,z) vectors that are orthogonal to the hyperplane spanned by $\mathrm{Re}\big((A_1)_{j,j+1},(A_1)_{j,j+2},(A_1)_{j+1,j+2}\big)$ and $\mathrm{Re}\big((A_2)_{j,j+1},(A_2)_{j,j+2},(A_2)_{j+1,j+2}\big)$ (except for the trivial degenerative case) are lying on a line of $\mathbb{R}^3$ crossing the centre $(0,0,0)$. Hence, let us use the parametrization $(\lambda x, \lambda y, \lambda z)$, where $\lambda \in \mathbb{R}$ and $x^2 + y^2 + z^2 = 1$, for the solutions of the above linear system.

Next, we choose the value of λ to obtain positive D_1 and D_2 and guarantee the lower rank for them as well.

For the lower rank of D_1 and D_2, it is enough that the following equalities are satisfied:

$$(\det D_1^>)(\lambda) = d_{-1}d_0d_1 + 2\lambda^3 xyz - \lambda^2(x^2 d_1 + y^2 d_0 + z^2 d_{-1}) = 0,$$
$$(\det D_2^>)(\lambda) = d_{-1}d_0d_1 - 2\lambda^3 \left(\frac{p}{1-p}\right)^3 xyz - \lambda^2 \left(\frac{p}{1-p}\right)^2 (x^2 d_1 + y^2 d_0 + z^2 d_{-1}) = 0.$$

To get positive semi-definite D_1 and D_2, let us calculate the Hilbert–Schmidt norm $\|\cdot\|_2$ of $D_1^>$ and $D_2^>$. We recall that

$$\|D_i^>\|_2^2 = \sum_{j,k} (D_i^>)_{jk}^2 = \mathrm{Tr}\,(D_i^>)^2 = \sigma^2_{D_i^>,1} + \sigma^2_{D_i^>,2} + \sigma^2_{D_i^>,3},$$

where $\sigma_{D_i^>,j}$ denote the eigenvalues of $D_i^>$. Since $\det D_i^> = 0$, the matrix $D_i^>$ is positive semi-definite if (and only if) $\|D_i^>\|_2 \leqslant \operatorname{Tr} D_i^>$ are satisfied $(i = 1, 2)$. Hence we need

$$\|D_1^>\|_2^2 = d_{-1}^2 + d_0^2 + d_1^2 + 2\lambda^2(x^2 + y^2 + z^2) \leqslant (\operatorname{Tr} D_1^>)^2,$$

which is the same as

(3.1)
$$|\lambda| \leqslant \sqrt{\frac{1}{2}\big((d_{-1} + d_0 + d_1)^2 - d_{-1}^2 - d_0^2 - d_1^2\big)} = (d_{-1}d_0 + d_0d_1 + d_1d_{-1})^{1/2}.$$

Using the analogous inequality for $\|D_2^>\|_2^2$, we note that it is enough to prove that the equation $(\det D_1^>)(\lambda) = 0$ has solutions of different signs λ_1, λ_2 such that (3.1) holds. Then one can find a $p \in (0, 1)$ such that $-\lambda_1 p/(1 - p) = \lambda_2$. Therefore, $(\det D_1^>)(\lambda_1) = (\det D_1^>)(\lambda_2) = (\det D_2^>)(\lambda_1) = 0$, which is what we intended to have. Moreover, the positivity of $D_i^>$ implies that $D_i \geqslant 0$, $i = 1, 2$.

To establish (3.1), note that the cubic function $\lambda \mapsto (\det D_1^>)(\lambda)$ $(\lambda \in \mathbb{R})$ has positive local maximum at zero, i.e. $(\det D_1^>)(0) = d_{-1}d_0d_1$. Hence we can find solutions of the equation $(\det D_1)(\lambda) = 0$ with the above property if (and only if)

$$(\det D_1^>)\big(\pm (d_{-1}d_0 + d_0d_1 + d_1d_{-1})^{1/2}\big) \leqslant 0$$

or, equivalently,

$$\begin{aligned} &d_{-1}d_0d_1 \pm 2(d_{-1}d_0 + d_0d_1 + d_1d_{-1})^{3/2}xyz \\ &\qquad - (d_{-1}d_0 + d_0d_1 + d_1d_{-1})(x^2d_1 + y^2d_0 + z^2d_{-1}) \leqslant 0. \end{aligned}$$

Expanding the last product we get

$$\begin{aligned} &\pm 2(d_1d_0 + d_0d_{-1} + d_{-1}d_1)^{3/2}xyz \\ &\qquad \leqslant x^2d_1^2(d_0 + d_{-1}) + y^2d_0^2(d_1 + d_{-1}) + z^2d_{-1}^2(d_1 + d_0). \end{aligned}$$

From the Cauchy–Schwarz inequality we have

$$\begin{aligned} &\big(d_1(d_0 + d_{-1}) + d_0(d_{-1} + d_1) + d_{-1}(d_1 + d_0)\big)xyz \\ &\qquad \leqslant \big(x^2d_1^2(d_0 + d_{-1}) + y^2d_0^2(d_1 + d_{-1}) + z^2d_{-1}^2(d_1 + d_0)\big)^{1/2} \\ &\qquad\quad \times \big(x^2y^2(d_1 + d_0) + y^2z^2(d_0 + d_{-1}) + x^2z^2(d_1 + d_{-1})\big)^{1/2}. \end{aligned}$$

Thus it is enough to prove that

(3.2)
$$\begin{aligned} &(d_1d_0 + d_0d_{-1} + d_1d_{-1})\big(x^2y^2(d_1 + d_0) + y^2z^2(d_0 + d_{-1}) + x^2z^2(d_1 + d_{-1})\big) \\ &\qquad \leqslant x^2d_1^2(d_0 + d_{-1}) + y^2d_0^2(d_1 + d_{-1}) + z^2d_{-1}^2(d_1 + d_0). \end{aligned}$$

After multiplication we see that the left-hand side is equal to

$$x^2(y^2+z^2)d_1^2(d_0+d_{-1})+y^2(x^2+z^2)d_0^2(d_1+d_{-1})$$
$$+z^2(x^2+y^2)d_{-1}^2(d_1+d_0)+2(x^2y^2+y^2z^2+z^2x^2)d_1d_0d_{-1}.$$

Since $x^2+y^2+z^2=1$, we can actually write (3.2) in the following form:

$$2(x^2y^2+y^2z^2+z^2x^2)d_1d_0d_{-1}$$
$$\leqslant x^4d_1^2(d_0+d_{-1})+y^4d_0^2(d_1+d_{-1})+z^4d_{-1}^2(d_1+d_0).$$

From the trivial identity $2(x^2y^2+y^2z^2+z^2x^2)=1-(x^4+y^4+z^4)$ it follows that

$$d_1d_0d_{-1}\leqslant(d_1d_0+d_0d_{-1}+d_{-1}d_1)(x^4d_1+y^4d_0+z^4d_{-1}),$$

which is that same as

$$\frac{1}{\frac{1}{d_1}+\frac{1}{d_0}+\frac{1}{d_{-1}}}=\frac{1}{\frac{x^2}{x^2d_1}+\frac{y^2}{y^2d_0}+\frac{z^2}{z^2d_{-1}}}\leqslant x^4d_1+y^4d_0+z^4d_{-1},$$

a weighted form of the harmonic and arithmetic mean inequality [7]. We showed that (3.2) holds, hence (3.1) also follows. This means that there exist $p\in(0,1)$ and real density matrices D_1,D_2 such that $D=pD_1+(1-p)D_2$, $\operatorname{Tr}D_iA_1=\operatorname{Tr}D_iA_2=0$ and $\operatorname{rank}D_i<\operatorname{rank}D$. ∎

Now we can prove our main result.

THEOREM 3.1. *Let $D\in M_n(\mathbb{C})$ be a density matrix and let $A_1,A_2\in M_n(\mathbb{C})$ be self-adjoint matrices. There exist a probability distribution p_i and a family of projections P_i such that*

$$D=\sum_i p_iP_i \quad \text{and} \quad \operatorname{Var}_D(A_1,A_2)=\sum_i p_i\operatorname{Var}_{P_i}(A_1,A_2).$$

Proof. Since

$$\operatorname{Var}_D(A_1,A_2)=\operatorname{Var}_D(A_1-\lambda_1I,A_2-\lambda_2I)$$

for any $\lambda_1,\lambda_2\in\mathbb{R}$, we can assume that $\operatorname{Tr}DA_1=\operatorname{Tr}DA_2=0$.

Let us apply induction on the rank of D. We remark that from conjugation by unitaries we can always assume that D is diagonal. If $\operatorname{rank}D=2$ then the existence of the decomposition is proved in Lemma 2.1.

For the general case, let us note that we can always reduce $\operatorname{rank}D$ according to Lemma 3.1. Since $\operatorname{Tr}D_iA_1=\operatorname{Tr}D_iA_2=0$ also holds, we readily get

$$\operatorname{Var}_D(A_1,A_2)=p\operatorname{Var}_{D_1}(A_1,A_2)+(1-p)\operatorname{Var}_{D_2}(A_1,A_2),$$

where $p\in(0,1)$. Now the induction gives a decomposition for D_1 and D_2 and the proof is complete. ∎

We remark here that the existence of the decomposition of the matrix variance is essentially based on the decomposition of the low-dimensional densities. In fact, the Krein–Milman theorem always gives the decomposition into rank-3 or lower rank densities. This is the subject of the next example. The proof is based on a simple geometrical observation.

LEMMA 3.2. *Let $\mathcal{R}$ be a nonempty intersection of an $(n-1)$-simplex Δ and two affine hyperplanes of $\mathbb{R}^n$ $(n \geqslant 3)$. The extreme points of $\mathcal{R}$ are lying at 2-simplices of Δ.*

Proof. If $n = 3$, we are ready. Thus let us assume that $n > 3$. An extreme point e of $\mathcal{R}$ must be lying on its topological boundary but any boundary point is on a proper face $\mathcal{F}_e$ of Δ. Since $\mathcal{F}_e \cap \mathcal{R}$ is a face of $\mathcal{R}$, it follows that e is an extreme point of $\mathcal{F}_e \cap \mathcal{R}$ as well. The simplex $\mathcal{F}_e$ has lower dimension, hence one can infer by induction that e is lying on a 2-simplex of Δ. ■

EXAMPLE 3.1. Let $n \geqslant 3$ and $D = \mathrm{Diag}(d_1, \dots, d_n) \in M_n(\mathbb{C})$ be a density matrix. Let A_1, $A_2 \in M_n(\mathbb{C})$ be self-adjoint matrices. Then one can find a probability distribution p_i and a family of densities D_i such that $\mathrm{rank} D_i \leqslant 3$,

$$D = \sum_i p_i D_i \quad \text{and} \quad \mathrm{Var}_D(A_1, A_2) = \sum_i p_i \mathrm{Var}_{D_i}(A_1, A_2).$$

In fact, again by the equality

$$\mathrm{Var}_D(A_1, A_2) = \mathrm{Var}_D(A_1 - \lambda_1 I, A_2 - \lambda_2 I),$$

we can assume that $\mathrm{Tr}\, DA_1 = \mathrm{Tr}\, DA_2 = 0$. Let Δ_{n-1} denote the convex hull of the standard basis vectors in $\mathbb{R}^n$, i.e. the standard $(n-1)$-simplex. The points $(r_1, \dots, r_n) \in \Delta_{n-1}$ that satisfy the equalities

$$\sum_i r_i (A_1)_{ii} = 0 \quad \text{and} \quad \sum_i r_i (A_2)_{ii} = 0$$

form a nonempty compact convex set $\mathcal{R}$ of $\mathbb{R}^n$. By Lemma 3.2, the extreme points of $\mathcal{R}$, denoted by $\mathcal{E}(\mathcal{R})$, are lying at 2-simplices of Δ_{n-1}. However, any element of $\mathcal{E}(\mathcal{R})$ is spanned by at most three standard basis vectors. On the other hand, the Krein–Milman theorem [4] gives

$$\mathcal{R} = \sum_{e \in \mathcal{E}(\mathcal{R})} p_e e, \quad \text{where} \sum_{e \in \mathcal{E}(\mathcal{R})} p_e = 1 \text{ and } p_e \geqslant 0.$$

Since any point in $\mathcal{E}(\mathcal{R})$ uniquely determines a diagonal operator (state) and $(d_1, \dots, d_n) \in \mathcal{R}$, we conclude that there exist densities D_i such that

$$D = \sum_i p_i D_i,$$

where $\operatorname{rank} D_i \leqslant 3$. Moreover, by the orthogonality $\operatorname{Tr} D_i A_1 = \operatorname{Tr} D_i A_2 = 0$ we obtain

$$\operatorname{Var}_D(A_1, A_2) = \sum_i p_i \operatorname{Var}_{D_i}(A_1, A_2).$$

We note that one can give an alternative proof of Theorem 3.1 relying on the previous example. In fact, any rank-3 density can be decomposed into lower rank ones according to the proof of Lemma 3.1 and Lemma 2.1. Furthermore, a linear estimate readily results from the number of the projections used in Theorem 3.1. Actually, by an elementary geometrical reasoning, the length of the decomposition is $O\big((\operatorname{rank} D)^3\big)$.

Acknowledgments. This work was partially supported by the Hungarian Research Grant OTKA K104206.

REFERENCES

[1] R. Bhatia, *Positive Definite Matrices*, Princeton University Press, Princeton 2007.
[2] O. Gühne, P. Hyllus, O. Gittsovich, and J. Eisert, *Covariance matrices and the separability problem*, Phys. Rev. Lett. 99 (2007), 130504.
[3] F. Hiai and D. Petz, *Introduction to Matrix Analysis and Applications*, to appear in Hindustan Book Agency.
[4] P. D. Lax, *Functional Analysis*, Wiley, 2002.
[5] D. Petz, *Quantum Information Theory and Quantum Statistics*, Springer, Heidelberg 2008.
[6] D. Petz and G. Tóth, *Matrix variances with projections*, Acta Sci. Math. (Szeged) 78 (2012), pp. 683–688.
[7] J. M. Steele, *The Cauchy–Schwarz Master Class*, Cambridge University Press, Cambridge 2004.
[8] G. Tóth and D. Petz, *Extremal properties of the variance and the quantum Fisher information*, Phys. Rev. A 87 (2013), 032324.
[9] A. Uhlmann, *Roofs and convexity*, Entropy 12 (2010), pp. 1799–1832.
[10] S. Yu, *Quantum Fisher information as the convex roof of variance*, http://arxiv.org/abs/1302.5311.

Alfréd Rényi Institute of Mathematics
H-1364 Budapest, POB 127, Hungary
E-mail: leka.zoltan@renyi.mta.hu

Alfréd Rényi Institute of Mathematics
H-1364 Budapest, POB 127, Hungary
E-mail: petz@math.bme.hu

Received on 10.3.2013;
revised version on 1.4.2013