EXTREME POINTS IN CONVEX SETS OF SYMMETRIC MATRICES

BERNARD YCART

ABSTRACT. This paper deals with the following problem: What are the extreme points of a convex set K of $n \times n$ matrices, which is the intersection of the set S_n of symmetric matrices of nonnegative type, with another convex subset of symmetric matrices H? In the case where the facial structure of H is known, we expose a general method to determine the extreme points of K (Theorem 1). Then, we apply this method to the set of correlation matrices, characterizing its extreme points in Theorem 2, which is our main theorem. A corollary describes thoroughly the extreme points of rank 2.

0. Introduction. The problem of determining the extreme points of particular convex sets of matrices, linked to various problems of optimization, has taken an increasing importance through the last twenty years, especially in the two directions of Mathematical Programming (cf. e.g. Berman [6]) and Data Analysis in Statistics. In the latter some questions are raised about the choice of new metrics satisfying particular conditions, for Principal Components Analysis (see Croquette [8]).

Those questions can often be reduced to problems of optimization over a convex set of symmetric matrices of nonnegative type.

We shall denote by S_n the set of $n \times n$ symmetric matrices of nonnegative type: S_n is a convex cone in the vector space of symmetric $n \times n$ matrices, identified with $\mathbf{R}^{n(n+1)/2}$. S_n and its convex subsets have often been studied (cf. Barker [1], Baston [2], Baumert [3 and 4], Hall [10] and Ycart [12]).

We consider a convex subset K of S_n, the intersection of S_n with another convex subset of $\mathbf{R}^{n(n+1)/2}$, H, of which the facial structure is supposed to be known. Our problem is to determine the extreme points of K.

In the first part, we expose a general method based upon simple geometrical and algebraic tools, that leads principally to a test deciding whether or not a given matrix in K is an extreme point (Theorem 1). A variant of this method, adapted to convex cones, had already proved to be successful in the case of those matrices of S_n whose coefficients are nonnegative (cf. [12]).

In the second part, we apply the method to the set of matrices in S_n whose diagonal coefficients are all equal to 1. Those matrices are of great interest for both statisticians and probabilists, as they can be interpreted as correlation matrices of random vectors in $\mathbf{R}^n$ (cf. Burington [7]). Theorem 2, which is the main theorem in this paper, characterizes the extreme points of the convex set of $n \times n$ correlation matrices. As a corollary, we describe all the extreme points of rank 2.

1. General method.

1.1. *Faces of a convex set*. We recall some definitions and elementary properties as derived from Rockafellar [11].

DEFINITION. A face of a convex set C is a convex subset C' of C such that every closed line segment in C with a relative interior point in C' has both end points in C'.

One knows that the collection of all relative interiors of nonempty faces of C is a partition of C. This fact justifies the following definition:

DEFINITION. Let x be a point in a convex set C. We call face of x relative to C, noted $FC(x)$, the face of C containing x in its relative interior.

EXAMPLE. x is an extreme point of C iff $FC(x) = \{x\}$.

We shall principally use the following proposition:

PROPOSITION 1. *Let A and B be two convex sets in $\mathbf{R}^N$, let x be a point in $A \cap B$. The face of x relative to $A \cap B$ is the intersection of the faces of x relative to A and B.*

PROOF. $FA(x) \cap FB(x)$ is obviously a face of $A \cap B$. As the relative interior of the intersection of two convex sets in $\mathbf{R}^N$ is the intersection of both relative interiors, if this intersection is not empty, the result follows.

1.2. *Faces of S_n*. The facial structure of S_n is often described in the literature. The reader will find detailed proofs in [1]. We shall use the following results.

PROPOSITION 2. *If X is a matrix in S_n of rank r, then the dimension of $FS_n(X)$ is $r(r+1)/2$.*

PROOF. Cf. [1, p. 29, Corollary 3].

PROPOSITION 3. *Let X and Y be two matrices of S_n. X and Y have the same face relative to S_n iff X, Y and $X + Y$ have the same rank.*

PROOF. By definition, X and Y belong to $FS_n((X + Y)/2)$; hence $FS_n(X) = FS_n(Y) = FS_n((X + Y)/2)$ iff the three matrices have the same rank (cf. [1, Lemma 4, pp. 28–29]).

1.3. *Necessary condition*. Let us give at once a quasi obvious condition that often proves to be useful in the searching of extreme points of $K = S_n \cap H$.

PROPOSITION 4. *If X is an extreme point of K, then*

$$\dim FH(X) \leqslant n(n+1)/2 - r(r+1)/2.$$

PROOF. If X is an extreme point of K, its face relative to K has dimension 0. Then it is sufficient to notice that if A and B are two convex sets of $\mathbf{R}^N$, having two points of their relative interiors in common, then

$$\dim A + \dim B \geqslant N + 1 \Rightarrow \dim A \cap B \geqslant 1,$$

which is a direct consequence of the classical result: let F and G be two subspaces of a finite-dimensional vector space E. Then $\dim F + \dim G \geqslant \dim E + 1$ implies that $\dim F \cap G \geqslant 1$.

1.4. *Lagrange decomposition.* For our study we need an algebraic tool which is a decomposition of the matrices of S_n.

Let $X = (x_{i,j})$, a nonzero matrix of S_n, let $i_0 = \text{Inf}\{i/x_{i,i} \neq 0\}$, and let $A(X)$ be the column matrix $[x_{i_0, j}(x_{i_0, i_0})^{-1/2}]$, $j = 1, \dots, n$. If $'A(X)$ is the transpose of $A(X)$, $A(X)'A(X)$ is an $n \times n$ matrix of rank 1, the i_0th line of which coincides with the i_0th line of X, the lines of lower index being zero for both $A(X)'A(X)$ and X.

Let $\varphi(X) = X - A(X)'A(X)$.

PROPOSITION 5. *If X is of rank r, then $\varphi(X)$ is in S_n, with rank $r - 1$.*

PROOF. Immediate.

As a direct consequence of Proposition 5, the iterates $\varphi^0(X) = X$, $\varphi(X), \dots, \varphi^r(X)$ belong to S_n and $\varphi^j(X)$ has rank $r - j$ if $0 \leqslant j \leqslant r$.

Let us note that $A_1 = A(X)$, $A_j = A(\varphi^{j-1}(X))$, $j = 1, \dots, r$. This sequence $(A_1, \dots, A_r)$ is well defined and unique for a given X, and we can write $X = \sum_{k=1}^r A_j\,'A_j$, the A_j's being independent.

This decomposition, unique in the sense we have described here, is of course a matricial version of the classical Lagrange method of decomposition for a quadratic form (cf. Gantmacher [**9**, Chapter 10]). It will be referred to from now on as "Lagrange decomposition of X."

1.5. *Necessary and sufficient condition.* Let us sum up in a theorem the algorithm of a test deciding whether or not a given matrix of K is an extreme point of K.

THEOREM 1. *Let X be a matrix of $K = S_n \cap H$ of rank r, $X = \sum_{j=1}^r A_j\,'A_j$, its Lagrange decomposition. X is an extreme point of K if and only if, for any matrix Y of K the three properties,*
  (i) $FH(X) = FH(Y)$,
  (ii) *the rank of Y is r,*
  (iii) *if $Y = \sum_{j=1}^r B_j\,'B_j$ is the Lagrange decomposition of Y, then for all $j = 1, \dots, r$, $(B_j, A_1, \dots, A_r)$ is a linearly dependent family,*
imply, $\forall j$, $B_j = A_j$.

PROOF. A matrix X of K is an extreme point iff any matrix Y, such that $FK(Y) = FK(X)$, is equal to X.

Thanks to Proposition 1, the only thing to prove is that (ii) and (iii) are equivalent to $FS_n(X) = FS_n(Y)$. By Proposition 3, this reduces to prove that (iii) is equivalent to: the rank of $X + Y$ is r.

We use the following classical lemma, easy to prove by induction.

LEMMA. *Let Q be a quadratic form, sum of squares of linear forms: $Q = f_1^2 + \cdots + f_p^2$. Then the rank of Q is the rank of the family of linear forms $(f_1, \dots, f_p)$.*

Let Q_X, Q_Y and Q_{X+Y} be the quadratic forms corresponding to the matrices X, Y and $X + Y$, respectively.

We have

$$Q_X = f_1^2 + \cdots + f_r^2, f_j \text{ linear form of matrix } A_j,$$

$$Q_Y = g_1^2 + \cdots + g_r^2, g_j \text{ linear form of matrix } B_j,$$

$$Q_{X+Y} = f_1^2 + \cdots + f_r^2 + g_1^2 + \cdots + g_r^2.$$

The lemma implies that Q_{X+Y} (hence $X + Y$) has rank r iff, $\forall j$, g_j is a linear combination of the f_i's; hence the result.

2. Extreme points of the set of correlation matrices. As an application of §1, let K be the set of $n \times n$ symmetric matrices of nonnegative type, whose diagonal coefficients are all equal to 1. Then H is the affine subspace of $\mathbf{R}^{n(n+1)/2}$ of those $n \times n$ symmetric matrices whose diagonal coefficients are equal to 1, dim $H = n(n + 1)/2 - n$.

We have $\forall X \in H$, $FH(X) = H$. Then Proposition 6 is an immediate application of Proposition 4.

PROPOSITION 6. *If X is an $n \times n$ matrix of K of rank r, the extreme point of K, then $r(r + 1)/2 \leqslant n$.*

EXAMPLE. For $n \leqslant 5$ (resp. 9), the extreme points of K have rank 1 or 2 (resp. 1, 2 or 3).

REMARK. The matrices of rank 1, belonging to extreme rays of the cone S_n, are necessarily extreme points of K.

If X of rank 1 is in K, there is an $n \times 1$ matrix $A = (a_i)$, $i = 1, \ldots, n$, such that

$$X = (a_i a_j)(i, j) = 1, \ldots, n, \qquad a_i^2 = 1 \ \forall i.$$

There are 2^{n-1} such matrices in K.

THEOREM 2. *Let X be an $n \times n$ matrix of rank r in K, $X = \sum_{j=1}^r A_j {}^t A_j$, its Lagrange decomposition, with ${}^t A_j = (a_{1,j}, \ldots, a_{n,j})$. X is an extreme point of K if and only if the identity matrix of order r is the only invertible matrix $(s_{i,j})$ in S_r verifying the system*

$$\sum_{i,j} s_{i,j} a_{k,i} a_{k,j} = 1, \qquad k = 1, \ldots, n.$$

(*It is a system in $r(r + 1)/2$ variables $s_{i,j}$, with n equations.*)

Geometrical interpretation. It is equivalent to say that X is an extreme point iff the n points of the projective space $\mathbf{P}(\mathbf{R}^r)$ whose projective coordinates are $(a_{k,1}, \ldots, a_{k,r})$, $k = 1, \ldots, n$, are not on the same nontrivial projective quadric in $\mathbf{P}(\mathbf{R}^r)$.

PROOF. We follow the test given in Theorem 1. Let Y be an $n \times n$ matrix of K with rank r and $Y = \sum_{j=1}^n B_j {}^t B_j$, its Lagrange decomposition. Let us suppose that for all j the family $(B_j, A_1, \ldots, A_r)$ is dependent. If A (resp. B) is the $n \times r$ matrix with columns A_j (resp. B_j), $j = 1, \ldots, r$, there exists an invertible $r \times r$ matrix M such that ${}^t B = M {}^t A$.

Let L_i (resp. N_i), $i = 1, \ldots, n$, be the column vectors such that ${}^t L_i = (a_{i,1}, \ldots, a_{i,r})$ (resp. ${}^t N_i = (b_{i,1}, \ldots, b_{i,r})$). X (resp. Y) is the Gram matrix (matrix of scalar products) of the L_i's (resp. N_i's).

$$X \text{ (resp. } Y) \in K \Leftrightarrow {}^t L_i L_i = 1 \text{ (resp. } {}^t N_i N_i = 1) \quad \forall i.$$

Now $\forall i = 1, \ldots, n$, $N_i = M L_i$. Hence we have

$$(1) \qquad\qquad {}^t N_i N_i = 1 = {}^t L_i {}^t M M L_i \quad \forall i = 1, \ldots, n.$$

Let us call $S = {}^t M M = (s_{i,j})(i, j) = 1, \ldots, r$: (1) is our system. Then, two cases are possible:

1) The system admits only one solution; necessarily $S = I$. Then M is an isometry of $\mathbf{R}^r$. X and Y, being the Gram matrices of two isometric systems of vectors, are equal; by Theorem 1, X is an extreme point.

2) The system admits a solution $S \neq I$, invertible and an element of S_r. Let M be a symmetric square root of S (see Gantmacher [9, Chapter 8]). Pose $N_i = M L_i$, $i = 1, \ldots, n$. Let Y be the Gram matrix of the N_i's; Y satisfies conditions (i)–(iii) of Theorem 1, but $Y \neq X$: X is not an extreme point.

COROLLARY. *The extreme points of K of rank 2 are the matrices $X = (x_{i,j})$ such that there exists $(\alpha_1, \ldots, \alpha_n) \in [0, 2\pi[^n$ verifying that three of the α_i's are different two by two modulo π, and $x_{i,j} = \cos(\alpha_i - \alpha_j), \forall i, j = 1, \ldots, n$.*

PROOF. Let $X = (x_{i,j})$ be an $n \times n$ matrix of K with rank 2, and $X = A_1 {}^t A_1 + A_2 {}^t A_2$ its Lagrange decomposition. We have $x_{i,j} = a_{i,1} a_{j,1} + a_{i,2} a_{j,2}$, and $\forall i$, $a_{i,1}^2 + a_{i,2}^2 = 1$. So, there exists $(\alpha_1, \ldots, \alpha_n) \in [0, 2\pi[^n$ such that, $\forall i$, $a_{i,1} = \cos \alpha_i$, $a_{i,2} = \sin \alpha_i$; hence $x_{i,j} = \cos(\alpha_i - \alpha_j)$.

A projective quadric in $\mathbf{P}(\mathbf{R}^2)$ is composed of one or two points. So, either the n projective points with coordinates $(\cos \alpha_i, \sin \alpha_i)$ vanish to only one or two different points, and X is not extremal, or not, and X is extremal.

REMARKS. (1) For $r = 3$, such a characterization is possible, although more tedious. The discussion is based upon the classical theorem: Five points of $\mathbf{P}(\mathbf{R}^3)$ such that any three of them are not on the same line determine a unique proper conic (cf. Berger [5]).

(2) In the general case, the problem of determining a projective quadric by a set of points in $\mathbf{P}(\mathbf{R}^r)$ is much more complicated, and not yet totally solved.

REFERENCES

1. G. P. Barker and D. Carlson, *Cones of diagonally dominant matrices*, Pacific J. Math. **57** (1975), 15–32.

2. V. J. D. Baston, *Extreme copositive quadratic forms*, Acta Arith. **15** (1969), 319–327.

3. L. D. Baumert, *Extreme copositive quadratic forms*, Pacific J. Math. **19** (1966), 197–204.

4. _____, *Extreme copositive quadratic forms*. II, Pacific J. Math. **20** (1967), 1–20.

5. M. Berger, *Géométrie T4*, CEDIC, Fernand Nathan, Paris, 1978.

6. A. Berman, *Cones, matrices and mathematical programming*, Lecture Notes in Econ. and Math., Springer-Verlag, Berlin, 1973.

7. R. S. Burington and D. C. May, *Handbook of probability and statistics* (2nd ed.), McGraw-Hill, New York, 1970.

8. A. Croquette, *Quelques résultats synthétiques en analyse des données multidimensionnelles: optimalité et métriques à effets relationnels*, Thèse 3ème cycle, Toulouse, 1980.

9. F. R. Gantmacher, *The theory of matrices*, Vol. I, Chelsea, New York, 1959.

10. M. Hall, Jr. and M. Newman, *Copositive and completely positive quadratic forms*, Proc. Cambridge Philos. Soc. **59** (1963), 329–339.

11. R. T. Rockafellar, *Convex analysis*, Princeton Math. Series, No. 28, Princeton Univ. Press, Princeton, N.J., 1970.

12. B. Ycart, *Extremales du cône des matrices de type non négatif, à coefficients positifs ou nuls*, Linear Algebra Appl. **48** (1982), 317–330.

Département de Mathématiques, U.P.P.A., Av. Louis Sallenave, 64000 Pau, France