

COVARIANCE SPACES FOR MEASURES ON POLYHEDRAL SETS

By J. H. B. KEMPERMAN¹ and MORRIS SKIBINSKY

Rutgers University and University of Massachusetts

Let V be a given subset of \mathbb{R}^n . We are interested in determining the associated moment space $\mathcal{C}_r[V]$. The latter consists of all points $c = (c(\mathbf{i}); |\mathbf{i}| \leq r)$ which can be realized as $c(\mathbf{i}) = \int x^{\mathbf{i}} \mu(dx)$, for all $|\mathbf{i}| \leq r$, by a measure μ on V . Here, $\mathbf{i} = (i_1, \dots, i_n)$ with $i_j \in \mathbb{Z}_+$ and $|\mathbf{i}| = i_1 + \dots + i_n$. Let $\mathcal{C}_r(V)$ be the analogous homogeneous moment space $\mathcal{C}_r(V)$, where one insists on $|\mathbf{i}| = r$. The calculation of $\mathcal{C}_r[V]$ is shown to be equivalent to that of $\mathcal{C}_r(W)$, with W as a suitable affine imbedding of V into \mathbb{R}^{n+1} . A central role is played by the dual $\mathcal{C}_r(V)^*$ of the convex cone $\mathcal{C}_r(V)$. One may interpret $\mathcal{C}_r(V)^*$ as the set of all homogeneous polynomials $f(x) = f(x_1, \dots, x_n)$ on \mathbb{R}^n of degree r that are nonnegative on V .

Detailed results are given only for the important case $r = 2$. Let \mathcal{Q}_n be the linear space of all symmetric $n \times n$ matrices, supplied with the natural inner product $(A, B) = \text{Tr}(AB)$. The pair $\mathcal{C}_2(V)$ and $\mathcal{C}_2(V)^*$ has a natural interpretation as a pair of dual convex cones in \mathcal{Q}_n . In fact, $\mathcal{C}_2(V)^*$ is the set of all $Q \in \mathcal{Q}_n$ such that $x^t Q x \geq 0$ for all $x \in V$. Special attention is given to the second order moment spaces $\mathcal{C}_2(K)$ and $\mathcal{C}_2[T]$ with

$$K = \{x \in \mathbb{R}^n : Ax \geq 0\}; \quad T = \{x \in \mathbb{R}^n : Bx + e \geq 0\}.$$

Here A and B denote given $m \times n$ matrices. Our description of the latter moment spaces involves the crucial cone $\mathcal{P}_m = \{Q \in \mathcal{Q}_m : x^t Q x \geq 0 \text{ for all } x \in \mathbb{R}_+^m\}$.

These results are quite explicit when $m \leq 4$, as happens, for instance, when T is a triangle in \mathbb{R}^2 or a simplex in \mathbb{R}^3 . This is largely due to the very simple structure of the cone \mathcal{P}_m in the case $m \leq 4$, due to Diananda (1962). The remaining problem, of determining the second order moment spaces $\mathcal{C}_2(K)$ or $\mathcal{C}_2[T]$ for the case $m \geq 5$, is essentially equivalent to the long standing difficult open problem to determine the precise structure of the cone \mathcal{P}_m when $m \geq 5$. Concrete applications will be given in subsequent papers.

¹Research supported in part by National Science Foundation Grant DMS-9002856. AMS 1991 subject classifications. 44A60, 60E15, 15A48, 49A55.

Key words and phrases. Finite multivariate moment problem, second order moment space, copositive matrices, nonnegative polynomials, dual cones, covariance spaces.

1. Moment Spaces

In the sequel, n and r are fixed positive integers. All sets considered are assumed to be measurable. All measures are finite-valued and nonnegative. For V as any subset of \mathbb{R}^n , let $\mathcal{M}_0(V)$ be the set of all measures μ on \mathbb{R}^n that are supported by a finite subset of V . The larger set of all measures μ supported by V and possessing all moments of order $\leq r$ will be denoted as $\mathcal{M}(V)$. Thus $\mu \in \mathcal{M}(V)$ if μ is a measure on V such that the integral

$$(1.1) \quad c(\mathbf{i}) = c_\mu(\mathbf{i}) = \int x^{\mathbf{i}} d\mu = \int x_1^{i_1} \cdots x_n^{i_n} \mu(dx)$$

is absolutely convergent for all $\mathbf{i} = (i_1, \dots, i_n) \in \mathbf{Z}_+^n$ with $|\mathbf{i}| = i_1 + \cdots + i_n \leq r$. Let

$$c[\mu] = (c_\mu(\mathbf{i}) : \mathbf{i} \in \mathbf{Z}_+^n; |\mathbf{i}| \leq r)$$

be the corresponding moment point (of order r); it has $\binom{n+r}{r}$ components. Let further

$$c(\mu) = (c_\mu(\mathbf{i}) : \mathbf{i} \in \mathbf{Z}_+^n; |\mathbf{i}| = r),$$

be the analogous “homogeneous” moment point (of order r); it has $\binom{n+r-1}{r}$ components.

We like to determine the set of all possible moment points $c[\mu]$ or $c(\mu)$, keeping V and r fixed. There is a considerable literature, see for instance Berg (1987), Cassier (1984) and Maserick (1977), on the analogous problem where $r = \infty$. Here, one is interested in characterizing the set of all *infinite* moment sequences $\{c_\mu(\mathbf{i}) : |\mathbf{i}| < \infty\}$ associated to the different measures μ on V . On the other hand, except for the very classical case $n = 1$, not much seems to be known about the *finite* case $2 \leq r < \infty$ we are considering.

Since only finitely many moments (1.1) are involved, we have for all $\mu \in \mathcal{M}(V)$ that there always exists a measure μ' on V of finite support ($\mu' \in \mathcal{M}_0(V)$) such that $c[\mu'] = c[\mu]$, thus also $c(\mu') = c(\mu)$. We will be interested in the two moment spaces

$$(1.2) \quad \begin{aligned} \mathcal{C}_r[V] &= \{c[\mu] : \mu \in \mathcal{M}(V)\} = \{c[\mu] : \mu \in \mathcal{M}_0(V)\} \quad \text{and} \\ \mathcal{C}_r(V) &= \{c(\mu) : \mu \in \mathcal{M}(V)\} = \{c(\mu) : \mu \in \mathcal{M}_0(V)\}. \end{aligned}$$

The right hand form (1.2) for $\mathcal{C}_r(V)$ essentially says that $\mathcal{C}_r(V)$ coincides with the convex cone spanned by the set of moments points $\{c(\delta_x) : x \in V\}$, one for each $x \in V$. Here, δ_x denotes the probability measure carried by $\{x\}$ thus

$$(1.3) \quad c(\delta_x) = (c(\mathbf{i}) = x_1^{i_1} \cdots x_n^{i_n} : \mathbf{i} \in \mathbf{Z}_+^n; |\mathbf{i}| = r).$$

Similarly for $\mathcal{C}_r[V]$. By homogeneity, $c(\delta_{\lambda x}) = \lambda^r c(\delta_x)$ for all $x \in \mathbb{R}^n$; $\lambda \geq 0$. Consequently, $\mathcal{C}_r(V)$ will remain unchanged when the subset V of \mathbb{R}^n is

modified by replacing each $x \in V$ by any non-empty subset (such as a single point) of the corresponding half ray $\{\lambda x : \lambda > 0\}$, and also when one adds or deletes the element $x = 0$. The resulting subset W of \mathbb{R}^n will be said to be equivalent to V , (since then $C_r(W) = C_r(V)$). In other words, V and W are *equivalent* when $\{\lambda x : \lambda > 0\} \cap V$ is non-empty if and only if $\{\lambda x : \lambda > 0\} \cap W$ is non-empty, this for each $x \in \mathbb{R}^n$ with $x \neq 0$. The convex cone W generated by V is equivalent to V , provided V is convex.

In many applications, the convex cone $K = \text{cone}(V)$ in \mathbb{R}^n generated by V happens to be a pointed cone. Equivalently, there are numbers ρ_1, \dots, ρ_n such that $\rho_1 x_1 + \dots + \rho_n x_n > 0$ for all $x \in V$ with $x \neq 0$. In that case, V is clearly equivalent to

$$(1.4) \quad W = \{x \in \mathbb{R}^n : \rho_1 x_1 + \dots + \rho_n x_n = 1; \lambda x \in V \text{ for some } \lambda > 0\},$$

in particular, $C_r(W) = C_r(V)$.

Let W be a subset of a hyperplane H in \mathbb{R}^n of the form $\sum_j \rho_j x_j = 1$ and consider any measure $\nu \in \mathcal{M}(W)$. We assert that the “non-homogeneous” moment point $c[\nu]$ is then already determined by the corresponding “homogeneous” moment point $c(\nu)$. In fact, any moment $c_\nu(\mathbf{i})$ of order $|\mathbf{i}| \leq r$ as in (1.1) can be expressed as an explicit linear combinations of the moments $c_\nu(\mathbf{k})$ of order $|\mathbf{k}| = r$. After all, expanding

$$x^\mathbf{i} = x_1^{i_1} \dots x_n^{i_n} = x_1^{i_1} \dots x_n^{i_n} [\rho_1 x_1 + \dots + \rho_n x_n]^{r-|\mathbf{i}|}, \quad (\text{for all } x \in W \subset H),$$

an integration relative to the measure ν yields that

$$(1.5) \quad c_\nu(\mathbf{i}) = \sum_{|\mathbf{j}|=r-|\mathbf{i}|} \binom{r-|\mathbf{i}|}{\mathbf{j}} \rho^\mathbf{j} c_\nu(\mathbf{i} + \mathbf{j}), \quad \text{whenever } |\mathbf{i}| \leq r.$$

In particular, (1.5) sets up a 1:1 correspondence between $C_r[W]$ and $C_r(W)$, and it only remains to determine the “homogeneous” moment space $C_r(W)$.

For T as an arbitrary subset of \mathbb{R}^d , one can reduce the study of the (non-homogeneous) moment space $C_r[T]$ to a study of an associated homogeneous moment space $C_r(W)$. Namely, take $W = \sigma T$ where $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^{d+1}$ is of the form

$$(1.6) \quad \sigma \mathbf{y} = (y_1, \dots, y_d, x_n) \text{ where } x_n = \frac{1}{\rho_n} (1 - \rho_1 y_1 - \dots - \rho_d y_d).$$

Here $n = d + 1$ and $\rho_n \neq 0$. Thus σ is a 1:1 affine map of \mathbb{R}^{n-1} onto the hyperplane $H = \{x \in \mathbb{R}^n : \sum_j \rho_j x_j = 1\}$. For example, if $H = \{x \in \mathbb{R}^n : x_n = 1\}$ then $\sigma \mathbf{y} = (y_1, \dots, y_{n-1}, 1)$. In all cases, $\pi = \sigma^{-1} : H \rightarrow \mathbb{R}^{n-1}$ is given by $\pi x = (x_1, \dots, x_{n-1})$.

For μ as a measure on \mathbb{R}^d , one has $\mu \in \mathcal{M}(T)$ if and only if $\sigma \mu \in \mathcal{M}(W)$. Here $\nu = \sigma \mu$ is defined by $\nu(B) = \mu(\sigma^{-1} B) = \mu(\pi B)$, for all $B \subset H$.

Moreover, one has $c_\mu(i_1, \dots, i_d) = c_\nu(i_1, \dots, i_d, 0)$ for all $\mathbf{i} = (i_1, \dots, i_d) \in \mathbf{Z}_+^d$ with $|\mathbf{i}| \leq r$, showing that $\mathcal{C}_r[T]$ may be regarded as a simple image of $\mathcal{C}_r[W]$. This allows us to reduce the study of $\mathcal{C}_r[T]$ to that of $\mathcal{C}_r[W]$, and thus, by (1.5), to a study of $\mathcal{C}_r(W)$.

2. Duality: The Case $r = 2$

A typical “homogeneous” moment point $c = (c(\mathbf{i}) : |\mathbf{i}| = r)$ will be regarded as a point $c \in \mathbb{R}^{n_0}$, where $n_0 = \binom{n+r-1}{r}$. Here, \mathbf{i} runs through all the n_0 tuples $\mathbf{i} = (i_1, \dots, i_n) \in \mathbf{Z}_+^n$ with $|\mathbf{i}| = r$. The inner product (α, c) in \mathbb{R}^{n_0} is given by $(\alpha, c) = \sum_{|\mathbf{i}|=r} \alpha(\mathbf{i})c(\mathbf{i})$.

Let $V \subset \mathbb{R}^n$. Recall that the homogeneous moment space $\mathcal{C}_r(V)$ is the convex cone generated by all points $c_x = c(\delta_x)$, one for each $x \in V$, (c_x having components $c_x(\mathbf{i}) = x^{\mathbf{i}}$). Thus, the dual of $\mathcal{C}_r(V)$ is the closed and convex cone given by

$$\begin{aligned} \mathcal{C}_r(V)^* &= \{ \alpha \in \mathbb{R}^{n_0} : (\alpha, c) \geq 0 \text{ for all } c \in \mathcal{C}_r(V) \} \\ &= \{ \alpha \in \mathbb{R}^{n_0} : (\alpha, c_x) \geq 0 \text{ for all } x \in V \}. \end{aligned}$$

Equivalently,

$$(2.1) \quad \mathcal{C}_r(V)^* = \{ \alpha \in \mathbb{R}^{n_0} : f_\alpha(x) \geq 0 \text{ for all } x \in V \}.$$

Here, $f_\alpha(x)$ denotes the homogeneous r^{th} degree polynomial

$$(2.2) \quad f_\alpha(x) = \sum_{|\mathbf{i}|=r} \alpha(\mathbf{i})x^{\mathbf{i}}, \quad (x \in \mathbb{R}^n; \alpha \in \mathbb{R}^{n_0}).$$

As is well known, the second dual $(\mathcal{C}_r(V)^*)^*$ is precisely the closure of the original convex cone $\mathcal{C}_r(V)$. Thus

$$(2.3) \quad \text{cl}(\mathcal{C}_r(V)) = \{ c \in \mathbb{R}^{n_0} : (\alpha, c) \geq 0 \text{ for all } \alpha \in \mathcal{C}_r(V)^* \}.$$

In many applications, $\mathcal{C}_r(V)$ is already closed. This is true for instance when V is compact and, hence, also when V is equivalent (as defined in Section 1) to a compact subset W of \mathbb{R}^n . Formula (2.3) remains true when $\mathcal{C}_r(V)^*$ is replaced by a subset \mathcal{E} of $\mathcal{C}_r(V)^*$, provided the convex cone generated by \mathcal{E} is dense in $\mathcal{C}_r(V)^*$. Often the latter property holds for the set $\mathcal{E}_r(V)$ of all extreme members α of $\mathcal{C}_r(V)^*$ and then (2.3) implies that

$$(2.4) \quad \text{cl}(\mathcal{C}_r(V)) = \{ c \in \mathbb{R}^{n_0} : (\alpha, c) \geq 0 \text{ for all } \alpha \in \mathcal{E}_r(V) \}.$$

Formula (2.3) is our starting point. It reduces the problem of determining the moment space $\mathcal{C}_r(V)$ (or rather its closure) to the problem of determining

all polynomials f_α as in (2.2) that are nonnegative on V . In this connection, see especially the interesting recent work of Micchelli and Pinkus (1989), which at least in spirit is close to the present paper.

Often (2.4) holds and then it suffices to determine all the extreme members of the collection of polynomials f_α as in (2.2) that are nonnegative on V . These problems tend to be very difficult. In a separate paper, we will apply the above ideas to the cubic case $r = 3$, in particular to the determination of $C_3[T]$ when T is a planar triangle.

From now on in the present paper, we assume that $r = 2$. In that case $n_0 = n(n + 1)/2$. Let \mathcal{Q}_n denote the linear space of all real and symmetric $n \times n$ matrices $Q = (q_{ij})$. Note that \mathcal{Q}_n has dimension n_0 . A second degree homogeneous polynomial $f_\alpha(x)$, as in (2.2) with $r = 2$, can be written as $f_\alpha(x) = x^t Q x$ with $Q = (q_{ij}) \in \mathcal{Q}_n$. Namely, let $\alpha(\mathbf{i}) = q_{11}$ if $\mathbf{i} = (2, 0, \dots, 0)$; $\alpha(\mathbf{i}) = 2q_{12} = 2q_{21}$ if $\mathbf{i} = (1, 1, 0, \dots, 0)$, and so on.

As to the moment point $c = (c(\mathbf{i}) : |\mathbf{i}| = 2) \in \mathbb{R}^{n_0}$, we prefer to represent it in the form of a matrix $C = (c_{ij}) \in \mathcal{Q}_n$. Namely, let $c(\mathbf{i}) = c_{11}$ if $\mathbf{i} = (2, 0, \dots, 0)$; $c(\mathbf{i}) = c_{12} = c_{21}$ if $\mathbf{i} = (1, 1, 0, \dots, 0)$ and so on. The original inner product (α, c) in \mathbb{R}^{n_0} now takes the form

$$(\alpha, c) = \sum_{|\mathbf{i}|=2} \alpha(\mathbf{i})c(\mathbf{i}) = \sum_{i=1}^n q_{ii}c_{ii} + \sum_{1 \leq i < j \leq n} 2q_{ij}c_{ij} = \sum_{i=1}^n \sum_{j=1}^n q_{ij}c_{ij} = \text{Tr}(QC),$$

which indeed is the natural inner product in the linear space \mathcal{Q}_n . The dual of a subset \mathcal{F} of \mathcal{Q}_n is defined as the closed convex cone $\mathcal{F}^* = \{A \in \mathcal{Q}_n : \text{Tr}(AB) \geq 0 \text{ for all } B \in \mathcal{F}\}$.

In the present case $r = 2$, the moment space $\mathcal{C}_2(V)$ can be identified with the convex cone in \mathcal{Q}_n defined by

$$(2.5) \quad \mathcal{C}_2(V) = \{C(\mu) : \mu \in \mathcal{M}(V)\} = \{C(\mu) : \mu \in \mathcal{M}_0(V)\}.$$

Here, $C(\mu) = C = (c_{ij}) \in \mathcal{Q}_n$ is given by

$$(2.6) \quad c_{ij} = \int x_i x_j \mu(dx), \quad (i, j = 1, \dots, n).$$

Note that $C \in S$ where $S = S_n$ will denote the closed and convex cone of all $Q \in \mathcal{Q}_n$ that are nonnegative definite, (also written as $Q \gg 0$).

PROPOSITION 1 *Let $C \in \mathcal{Q}_n$. Then $C \in \mathcal{C}_2(V)$ if and only if C can be written as $C = U^t U$ with $U = (u_{ij})$ as an $s \times n$ matrix, such that*

$$(2.7) \quad (u_{k1}, \dots, u_{kn}) \in \{\lambda x : x \in V; \lambda \geq 0\}, \quad (k = 1, \dots, s).$$

Here, s may depend on C . In the special case $V = \mathbb{R}_+^n$ this requires precisely that $C = U^t U$ for some $U \geq 0$.

PROOF From (2.5), $C \in \mathcal{C}_2(V)$ if and only if $C = C(\mu)$ for some measure μ having a finite support $\{x^{(1)}, \dots, x^{(s)}\} \subset V$. That is, $c_{ij} = \sum_k p_k x_i^{(k)} x_j^{(k)}$ for all $i, j = 1, \dots, n$, where $p_k = \mu(\{x^{(k)}\}) > 0$ ($k = 1, \dots, s$). Letting $u_{kj} = \sqrt{p_k} x_j^{(k)}$ one obtains the stated assertion.

REMARK As is clear from the proof, the different ways of writing C as $C = U^t U$ with U satisfying (2.7) correspond precisely to the different measures μ on V of finite support such that $C = C(\mu)$. The stated property is equivalent to $x^t C x$ being representable as a finite sum $x^t C x = \sum_{k=1}^s L_k(x)^2$ with the $L_k(x)$ as linear forms $L_k(x) = u_{k1} x_1 + \dots + u_{kn} x_n$ ($k = 1, \dots, s$) satisfying condition (2.7). In the special case $V = \mathbb{R}_+^n$, such matrices C are also said to be completely positive, see Hall (1986, p. 350).

Presently, (2.1) and (2.3) take the form

$$(2.8) \quad \text{cl}(\mathcal{C}_2(V)) = \{C \in \mathcal{Q}_n : \text{Tr}(QC) \geq 0 \text{ for all } Q \in \text{cop}(V)\} = \text{cop}(V)^*.$$

DEFINITION By $\text{cop}(V)$ we mean the closed and convex cone defined by

$$(2.9) \quad \text{cop}(V) = \mathcal{C}_2(V)^* = \{Q \in \mathcal{Q}_n : x^t Q x \geq 0 \text{ for all } x \in V\}.$$

The matrices $Q \in \text{cop}(V)$ are said to be V -copositive.

Note that (2.8) remains valid when $\text{cop}(V)$ is replaced by a subset \mathcal{E} of $\text{cop}(V)$ such that the convex cone generated by \mathcal{E} is dense in $\text{cop}(V)$. Our main task would be the explicit determination of such a set \mathcal{E} or if possible of the set $\text{cop}(V)$ itself.

NOTATION Let $S = S_n$ denote the class of nonnegative definite $Q \in \mathcal{Q}_n$. Let further $N = N_n$ denote the closed and convex cone consisting of all nonnegative $Q \in \mathcal{Q}_n$, (written as $Q \geq 0$). Let further $N + S$ denote the set of all sums $Q_1 + Q_2$ with $Q_1 \in N$ and $Q_2 \in S$. Note that $N + S$ is precisely the convex cone generated by $N \cup S$.

It is evident that $S \subset \text{cop}(V)$ for every $V \subset \mathbb{R}^n$. If $V \subset \mathbb{R}_+^n$, then also $N \subset \text{cop}(V)$ and thus $N + S \subset \text{cop}(V)$. The set $N + S$ can easily be much smaller than $\text{cop}(V)$. For instance, let $1 \leq d < n$ and suppose each $x \in V$ satisfies $x_j = 0$ for $d < j \leq n$, while $x_j \geq 0$ otherwise, (so that $V \subset \mathbb{R}_+^n$). In this case, the condition $Q = (q_{ij}) \in \text{cop}(V)$ clearly depends only on the elements q_{ij} with $1 \leq i, j \leq d$ and thus there is no need at all that $Q \in N + S$.

Of central importance is the case $V = \mathbb{R}_+^n$. We will write $\mathcal{P} = \mathcal{P}_n = \text{cop}(\mathbb{R}_+^n)$. Thus \mathcal{P} is the closed and convex cone consisting of all $Q \in \mathcal{Q}_n$ such that

$$(2.10) \quad x^t Q x \geq 0 \text{ for all } x \geq 0.$$

Such matrices Q are said to be copositive. Note that $N + S \subset \mathcal{P}$.

Observe that $V = \mathbb{R}_+^n$ is equivalent to the *compact* simplex

$$(2.11) \quad S(n) = \{x = (x_1, \dots, x_n) : x_i \geq 0 (i = 1, \dots, n); x_1 + \dots + x_n = 1\}.$$

This implies that $C_2(\mathbb{R}_+^n) = C_2(S(n))$ is closed, hence, (2.8) presently takes the form

$$(2.12) \quad C_2(S(n)) = \mathcal{P}^* = \{C \in \mathcal{Q}_n : \text{Tr}(QC) \geq 0 \text{ for all } Q \in \mathcal{P}\}.$$

Thus, if we would precisely know the class \mathcal{P} then we would also know what $C \in \mathcal{Q}_n$ can be realized as $C = C(\mu)$ by a measure on $S(n)$ and, as a consequence, also what sets of first and second moments can be realized by a measure on the simplex $T(n - 1)$ in \mathbb{R}^{n-1} . Here,

$$(2.13) \quad T(d) = \{y = (y_1, \dots, y_d) \in \mathbb{R}^d : y_i \geq 0 (i = 1, \dots, d); y_1 + \dots + y_d \leq 1\}$$

For arbitrary n , one has that $N + S \subset \mathcal{P}$ thus

$$(2.14) \quad C_2(S(n)) = \mathcal{P}^* \subset (N + S)^* = N \cap S.$$

After all $(N + S)^* = (N \cup S)^* = N^* \cap S^* = N \cap S$, since $N^* = N$ and $S^* = S$ as is easily seen, (see also Hall (1986, p. 353)). Here, $S^* = S$ says that $A \in \mathcal{Q}_n$ belongs to S if and only if $\text{Tr}(AB) \geq 0$ for all $B \in S$.

Unfortunately, a precise description of the class $\mathcal{P} = \text{cop}(\mathbb{R}_+^n)$ is only available when $n \leq 4$. In fact, Diananda (1962) showed that $\mathcal{P} = N + S$ when $n \leq 4$. Therefore,

$$(2.15) \quad C_2(S(n)) = \mathcal{P}^* = (N + S)^* = N \cap S \text{ if } n \leq 4.$$

If $n \geq 5$ then $N + S$ is a proper subset of \mathcal{P} . This follows from an unpublished counterexample due to A. Horn, which is discussed in Hall (1986, p. 357). Specifically, Horn gave an explicit example with $n = 5$ of a member $H \in \mathcal{P}/(N + S)$. It is given below, see (4.13). It does not seem to be known whether or not $N + S$ is closed. Anyway, assuming $n \geq 5$, we can show (see the last remark of the paper) that even the closure of $N + S$ is a proper subset of \mathcal{P} , equivalently \mathcal{P}^* is a proper subset of $N \cap S$. In this way, we arrive at the following result.

THEOREM 1 *In order that an $n \times n$ symmetric matrix $C = (c_{ij})$ admits a representation as in (2.6) with μ as a measure on the simplex $S(n)$ defined by (2.11), it is necessary that $C \in N \cap S$. Equivalently, C must be nonnegative definite and such that*

$$(2.16) \quad c_{ij} \geq 0 \text{ for all } 1 \leq i < j \leq n.$$

This necessary condition $C \in N \cap S$ is also sufficient when $n \leq 4$, but not when $n \geq 5$.

3. The Polyhedral Case

Here, we assume that the subset V of \mathbb{R}^n is a polyhedral cone of the form

$$(3.1) \quad K = \{x \in \mathbb{R}^n : Ax \geq 0\}.$$

Here, A is a given $m \times n$ real valued matrix. Recall that each $C = C(\mu)$ is necessarily nonnegative definite. A central result is as follows.

THEOREM 2 *Let $C \in \mathcal{Q}_n$ be nonnegative definite. Then in order that $C \in \text{cl}(\mathcal{C}_2(K))$ it is necessary and sufficient that*

$$(3.2) \quad \text{Tr}(A^t P A C) \geq 0 \text{ for all } P \in \mathcal{P}_m.$$

A necessary condition is that

$$(3.3) \quad A C A^t \geq 0$$

Conversely, if $m \leq 4$ then (3.3) is also sufficient for $C \in \text{cl}(\mathcal{C}_2(K))$.

REMARK Let $P \in \mathcal{P}_m$, that is, $P \in \mathcal{Q}_m$ and $y^t P y \geq 0$ for all $y \in \mathbb{R}_+^m$. Hence, $x^t A^t P A x \geq 0$ for all $x \in K$. Integrating the latter inequality relative to a measure on μ on K such that $C(\mu) = C$, this confirms that (3.2) is a necessary condition for $C \in \mathcal{C}_2(K)$ and thus for $C \in \text{cl}(\mathcal{C}_2(K))$. In practice, we are yet unable to verify the necessary and sufficient condition (3.2) when $m \geq 5$, simply because then the class \mathcal{P}_m is still largely unknown. The necessary condition (3.3) is equivalent to

$$(3.4) \quad \sum_{i=1}^n \sum_{j=1}^n a_{gi} a_{hj} c_{ij} \geq 0 \text{ for all } 1 \leq g < h \leq m,$$

altogether $\binom{m}{2}$ conditions. The necessity of (3.4) for $C \in \mathcal{C}_2(K)$ is obvious from an integration of the quadratic function $f(x) = (Ax)_g (Ax)_h$ relative to a measure μ on K such that $C(\mu) = C$. From (3.1), $f(x) \geq 0$ for all $x \in K$.

PROOF Suppose $P \in \mathcal{P}_m$ thus $y^t P y \geq 0$ when $y \in \mathbb{R}^m$, $y \geq 0$. Substituting $y = Ax$, we see that $x^t A^t P A x \geq 0$ for all $x \in K$, that is, $A^t P A \in \text{cop}(K)$ as defined by (2.9). It follows from (2.8) that (3.2) is a necessary condition.

Sufficiency. Suppose (3.2) holds. Let $\mathcal{G} = A^t \mathcal{P}_m A + S$ be the subset of $\text{cop}(K)$ which consists of all $Q \in \mathcal{Q}_n$ of the form $Q = A^t P A + B$ with $P \in \mathcal{P}_m$ and $B \in S$. We see from (3.2) and $C \in S$ that $\text{Tr}(Q C) \geq 0$ for all $Q \in \mathcal{G}$. In view of (2.8), this would imply $C \in \text{cl}(\mathcal{C}_2(K))$, provided \mathcal{G} can be shown to be dense in $\text{cop}(K)$.

In fact, let $Q \in \text{cop}(K)$ be given and consider $Q(\delta) = Q + \delta I$, with I as the $n \times n$ identity matrix and $\delta > 0$. It suffices to show that $Q(\delta) \in \mathcal{G}$ for all $\delta > 0$. In fact, merely using the fact that $x^t Q(\delta)x > 0$ whenever $Ax \geq 0$ and $x \neq 0$, this result $Q(\delta) \in \mathcal{G}$ is an immediate consequence of Theorem 4.2 due to Martin and Jacobson (1981).

By the way, Martin, Powell and Jacobson (1981, p. 53) showed by example that \mathcal{G} can be a proper subset of $\text{cop}(K)$. See also Martin and Jacobson (1981, p. 246).

We already saw that condition (3.3) is necessary for $C \in \mathcal{C}_2(K)$ and thus for $C \in \text{cl}(\mathcal{C}_2(K))$. Also observe that (3.3) is equivalent to $\text{Tr}(A^t BAC) \geq 0$ for all $B \in N_m$, that is, for all nonnegative $B \in \mathcal{Q}_m$. As to the sufficiency of (3.3), assuming $m \leq 4$, it suffices to show that (3.3) implies (3.2). In fact, Diananda (1962) showed, for $m \leq 4$, that each $P \in \mathcal{P}_m$ is of the form $P = B + Q$ where $B \in N_m$ and $Q \in S_m$. Therefore

$$\text{Tr}(A^t PAC) = \text{Tr}(A^t BAC) + \text{Tr}(A^t QAC) \geq 0.$$

Here, it is also used that $A^t QA \in S_n$ and $C \in S_n$.

COMMENTS Suppose K is a direct product $K = K_1 \times \mathbb{R}^{n-d}$ with K_1 as a polyhedral cone in \mathbb{R}^d . Equivalently, $a_{ij} = 0$ when $j > d$. It is interesting to note that then the necessary and sufficient condition (3.2) depends only on the c_{ij} with $1 \leq i, j \leq d$. This feature also follows from (2.8) and the following Proposition 2, (which is related to Lemma 9 in Diananda (1962)). Here, $\text{span}(V)$ denotes the linear span of V .

PROPOSITION 2 *Let the subset V of \mathbb{R}^n be a direct product $V = W \times \mathbb{R}^{n-d}$, with W as a subset of \mathbb{R}^d , ($1 \leq d < n$). Assume that $\text{span}(V) = \mathbb{R}^n$, equivalently, $\text{span}(W) = \mathbb{R}^d$. Then a matrix $P \in \mathcal{Q}_n$ belongs to $\text{cop}(V)$ if and only if there exist matrices $Q \in \text{cop}(W)$ and $B \in S_n$ such that $x^t Px = y^t Qy + x^t Bx$, for all $x \in \mathbb{R}^n$. Here, we write $x \in \mathbb{R}^n$ as $x = (y, z)$ with $y \in \mathbb{R}^d$ and $z \in \mathbb{R}^{n-d}$. Thus $Q \in \mathcal{Q}_m$ and $y^t Qy \geq 0$ if $y \in \mathbb{R}_+^d$.*

PROOF Using an induction with respect to $k = n - d$, it suffices to consider the case $d = n - 1$. The stated condition is clearly sufficient. Conversely, assume that $P \in \text{cop}(V)$, that is, $x^t Px \geq 0$ whenever $x \in V$, that is, whenever $y \in W$. Let G be the upper principal $d \times d$ submatrix of P and let $L(y) = a_{n1}y_1 + \dots + a_{n,n-1}y_{n-1}$. Then $y \in W$ implies that

$$(3.5) \quad x^t Px = a_{nn}z^2 + 2L(y)z + y^t Gy \geq 0 \text{ for all } z \in \mathbb{R}.$$

If $a_{nn} = 0$, it follows that $L(y) = 0$ for all $y \in W$, hence, $L(y) \equiv 0$ (since $\text{span}(W) = \mathbb{R}^d$) and the stated assertion holds with $Q = G$ and $B = 0$. In the case $a_{nn} > 0$ one may as well assume that $a_{nn} = 1$. Let $Q \in \mathcal{Q}_{n-1}$

be defined by $y^t Q y \equiv y^t G y - L(y)^2$. It follows from (3.5) that $y^t Q y \geq 0$, for all $y \in W$, hence, $Q \in \text{cop}(W)$. We further have for all $x \in \mathbb{R}^n$ that $x^t P x = y^t Q y + x^t B x$, where $x^t B x \equiv (L(y) + z)^2 \geq 0$ thus $B \in S_n$.

4. Applications

Detailed applications of the preceding theory will be given in a separate paper. Here we will sketch just one application. Let T be a polyhedral convex subset of \mathbb{R}^d of the form

$$(4.1) \quad T = \{y \in \mathbb{R}^d : B y + e \geq 0\}.$$

Here, B is an $m \times d$ matrix and e an $m \times 1$ column vector. One may as well assume that $0 \in T$, hence, $e \geq 0$ and further that T has at least two points. For convenience, we will further assume that T is compact thus $m \geq d + 1$; (the non-compact case can be handled equally well but is somewhat more delicate). Note that $e \neq 0$, (otherwise, T would be unbounded).

We like to determine the moment space $\mathcal{C}_2[T]$. Equivalently, we are interested in the necessary and sufficient conditions on the numbers c_0, ξ_i and c_{ij} ($i, j = 1, \dots, d$) in order that $\mu \in \mathcal{M}(T)$ can be found such that

$$(4.2) \quad \begin{aligned} c_0 &= \int \mu(dy); & \xi_i &= \int y_i \mu(dy); \\ c_{ij} &= \int y_i y_j \mu(dy), & (i, j &= 1, \dots, d). \end{aligned}$$

If $c_0 = 1$ then $\xi_i = E Y_i$ and $c_{ij} = E Y_i Y_j$ when (Y_1, \dots, Y_d) takes its values in T and has distribution μ . In terms of $X = (X_1, \dots, X_n) = (Y_1, \dots, Y_{n-1}, 1)$ one also has $c_0 = E X_n^2$ and $\xi_i = E X_i X_n$ ($1 \leq i < n$).

Let $n = d + 1$. Points $x \in \mathbb{R}^n$ will be written as $x = (y, x_n)$ where $y \in \mathbb{R}^{n-1}$ and $x_n \in \mathbb{R}$. It will be convenient to identify T with the compact set

$$(4.3) \quad \begin{aligned} V &= \{x = (y, x_n) \in \mathbb{R}^n : y \in T; x_n = 1\} \\ &= \{x \in \mathbb{R}^n : A x \geq 0; x_n = 1\}. \end{aligned}$$

Here, A denotes the $m \times n$ matrix $A = (B, e)$. Further let

$$c_{nn} = \xi_n = c_0 \quad \text{and} \quad c_{in} = c_{ni} = \xi_i \quad (1 \leq i \leq n).$$

The above question now reduces to the problem of determining the homogeneous moment space $\mathcal{C}_2(V)$ of all $C = (c_{ij}) \in \mathcal{Q}_n$ which can be realized as $c_{ij} = \int x_i x_j \nu(dx)$ ($1 \leq i, j \leq n$) by a finite measure ν on V . Since V is compact, the convex cone $\mathcal{C}_2(V)$ is closed.

Next consider the closed and convex cone

$$(4.4) \quad K = \{x \in \mathbb{R}^n : Ax \geq 0\}.$$

We claim that K is precisely the convex cone generated by V . Since V is convex this implies that K and V are equivalent, thus $\mathcal{C}_2(K) = \mathcal{C}_2(V)$.

By $V \subset K$ one has $\text{cone}(V) \subset K$. As to the converse, assume that $x = (y, x_n) \neq 0$ belongs to K thus $Ax \geq 0$. It suffices to show that $\lambda x \in V$ for some $\lambda > 0$, equivalently, that $x_n > 0$. On the contrary, suppose that $x_n \leq 0$. Since $Ax = By + bx_n \geq 0$ one has that $By \geq (-x_n)e \geq 0$ (since $e \geq 0$). Hence, $y = 0$ since, otherwise, T would be unbounded, (in view of $T + \beta y \subset T$ for all $\beta \geq 0$). Next $x = (0, x_n) \neq 0$, thus, $x_n < 0$. But now $0 = By \geq (-x_n)e \geq 0$ implies that $e = 0$ and we have a contradiction.

We can now apply Theorem 2. The condition that the $n \times n$ matrix $C = (c_{ij})$ be nonnegative definite is easily checked. One may as well assume that $c_0 > 0$. It is natural to introduce the quantities

$$(4.5) \quad \sigma_{ij} = c_{ij} - \xi_i \xi_j / c_0, \quad (i, j = 1, \dots, n).$$

Especially note that $\sigma_{in} = \sigma_{ni} = 0$ for all i . If $c_0 = 1$ then the σ_{ij} can be regarded as covariances $\sigma_{ij} = \text{Cov}(Y_i, Y_j)$ ($i, j = 1, \dots, n$), where $Y_n \equiv 1$. As is easily seen, the $n \times n$ matrix $C = (c_{ij})$ is nonnegative definite if and only if $\Sigma = (\sigma_{ij}; i, j = 1, \dots, d)$ is nonnegative definite.

It remains to check the necessary and sufficient condition (3.2) of Theorem 2. Also in view of the remark following Theorem 2, that condition amounts to requiring that, for each $P = (p_{rs}) \in \mathcal{P}_m$, the left hand side of the obvious inequality

$$(4.6) \quad \sum_{r=1}^m \sum_{s=1}^m p_{rs} \left[\sum_{i=1}^d b_{ri} y_i + e_r \right] \left[\sum_{j=1}^d b_{sj} y_j + e_s \right] \geq 0 \quad \text{for all } y \in T,$$

must integrate to a nonnegative number relative to any measure μ on T that satisfies (4.2). In particular, for each choice of $1 \leq r < s \leq m$, the same must be true for the inequality

$$(4.7) \quad \left[\sum_{i=1}^d b_{ri} y_i + e_r \right] \left[\sum_{j=1}^d b_{sj} y_j + e_s \right] \geq 0 \quad \text{for all } y \in T.$$

Naturally, (4.7) is merely a special case of (4.6) since $N \subset \mathcal{P}_m$.

For convenience we take $c_0 = 1$. Using (4.5), it is easily seen that (4.7) leads to the necessary condition, that for all $1 \leq r < s \leq m$,

$$(4.8) \quad \sum_{i=1}^d \sum_{j=1}^d b_{ri} b_{sj} \sigma_{ij} + \left[\sum_{i=1}^d b_{ri} \xi_i + e_r \right] \left[\sum_{j=1}^d b_{sj} \xi_j + e_s \right] \geq 0;$$

this also holds for $r = s$ because $\Sigma \gg 0$. Theorem 2 further implies that (4.8) is also sufficient (for μ to exist) provided $m \leq 4$. In particular, $\Sigma \gg 0$ together with (4.8) must imply that $\xi = (\xi_1, \dots, \xi_d) \in T$, at least when $m \leq 4$.

The following Theorem 3 summarizes some of the above results and contains Theorem 1 as a special case. Here, the ξ_i and $\sigma_{ij} = \sigma_{ji}$ ($i, j = 1, \dots, d$) are given numbers.

THEOREM 3 *Let T be the form (4.1). In order that there exist random variables Y_1, \dots, Y_d satisfying $\Pr((Y_1, \dots, Y_d) \in T) = 1$; $EY_i = \xi_i$ and $\text{Cov}(Y_i, Y_j) = \sigma_{ij}$ ($i, j = 1, \dots, d$) it is necessary that $\Sigma = (\sigma_{ij})$ be nonnegative definite and satisfies (4.8). If $m \leq 4$ then these necessary conditions are also sufficient.*

COUNTER EXAMPLE The following example shows that the above (rather obvious) necessary conditions on the ξ_j and c_{ij} are not sufficient anymore with $m \geq 5$. Choose $c_0 = 1$, $d = 4$ and T as the four-dimensional simplex $T = T(4)$, as in (2.13), consisting of all $y = (y_1, \dots, y_4) \in \mathbb{R}^4$ satisfying $y_i \geq 0$ ($i = 1, \dots, 5$). Here and below,

$$y_5 = 1 - (y_1 + y_2 + y_3 + y_4).$$

Thus, T is of the form (4.1) with $m = 5$. Namely, $b_{ij} = \delta_i^j$ if $i = 1, \dots, 4$; $b_{5,j} = -1$ and $e_j = \delta_j^5$, ($j = 1, \dots, 4$). The special system (4.7) (of functions nonnegative on T) now consists of the 10 functions $y_r y_s$ ($1 \leq r, s \leq 5$; $r < s$). Hence, the necessary conditions (4.8) take the form

$$(4.9) \quad c_{rs} \geq 0; \quad \bar{c}_{r5} \geq 0 \quad (1 \leq r < s \leq 4), \quad \text{where } \bar{c}_{r5} = \bar{c}_{5r} = \xi_r - \sum_{j=1}^4 c_{rj}.$$

Note that \bar{c}_{r5} corresponds to the integral of $y_r y_5$. Let us choose

$$(4.10) \quad c_0 = 1; \quad \xi_i = \frac{1}{5} \quad \text{and} \quad c_{ij} = d_{j-i} \quad (i, j = 1, \dots, 4),$$

with $d_j = d_{-j} = d_{j-5}$. Put $d_0 = \alpha$; $d_{\pm 1} = \beta$; $d_{\pm 2} = \gamma$. For instance, $c_{14} = d_3 = d_{-2} = d_2 = \gamma$. We will further assume that

$$(4.11) \quad \alpha > \beta > \gamma > 0 \quad \text{and} \quad \alpha + 2\beta + 2\gamma = \frac{1}{5},$$

(thus $\gamma < 1/25 < \alpha$). This implies that $\bar{c}_{r5} = d_{5-r} > 0$ ($1 \leq r \leq 4$), such as $\bar{c}_{15} = d_4 = d_{-1} = \beta$ and $\bar{c}_{45} = \beta$. The necessary conditions (4.9) are now automatically satisfied. We will further assume that the 4×4 matrix $\Sigma = (\sigma_{ij})$ is strictly positive definite. It is easily seen that this is true if and only if

$$(4.12) \quad \alpha(\alpha - \beta - \gamma) > \beta^2 - 3\beta\gamma + \gamma^2 \quad \alpha(\alpha + \gamma) > 2\beta^2.$$

The necessary conditions of Theorem 3 are now all satisfied. Recall that these conditions essentially derive, as in (3.2), from the different members $P \in N + S \subset \mathcal{P}_5$. However, there is no guarantee that the moment problem on hand does have a solution, precisely because \mathcal{P}_5 happens to be strictly larger than $N + S$. One must also insist that the ξ_j and c_{ij} satisfy all the moment conditions associated, as in (3.2), to the different members $H \in \mathcal{P}_5/(N + S)$. One such matrix H is defined by the so-called Horn form

$$(4.13) \quad \begin{aligned} y^t H y &= (y_1 + y_2 + y_3 + y_4 + y_5)^2 \\ &\quad - 4(y_1 y_2 + y_2 y_3 + y_3 y_4 + y_4 y_5 + y_5 y_1), \end{aligned}$$

(already mentioned at the end of Section 2). Since

$$\begin{aligned} y^t H y &= (y_1 - y_2 + y_3 + y_4 - y_5)^2 + 4y_2 y_4 + 4y_3(y_5 - y_4) \\ &= (y_1 - y_2 + y_3 - y_4 + y_5)^2 + 4y_2 y_5 + 4y_1(y_4 - y_5), \end{aligned}$$

from Hall (1986, p. 357), one has $y^t H y \geq 0$ for all $y \in \mathbb{R}_+^5$ thus $H \in \mathcal{P}_5$. Integrating $y^t H y$ relative to a probability measure μ on $T(4)$ satisfying (4.2) (with $c_0 = 1$), this special matrix H leads to the new necessary condition

$$(4.14) \quad c_{12} + c_{23} + c_{34} + \bar{c}_{45} + \bar{c}_{51} \leq \frac{1}{4};$$

(permuting indices leads to a set of $120/5 = 24$ different necessary conditions of type (4.14)). In the present example, (4.14) is equivalent to $5\beta \leq 1/4$. Thus we have the desired counter example as soon as $\beta > 1/20$. In fact, choose $0 < \delta < \delta_0 = 2/(11 + \sqrt{125}) = .09017$. Then the parameters

$$(4.15) \quad \gamma = 1/(80 + 10\delta); \quad \alpha = 6\gamma; \quad \beta = (4 + \delta)\gamma$$

do satisfy $\alpha > \beta > \gamma > 0$, further (4.11), (4.12) as well as $\beta > 1/20$.

REMARK By the way, substituting (4.15), the first inequality (4.12) is equivalent to $1 - 11\delta - \delta^2 > 0$ and becomes an equality when $\delta = \delta_0$, showing that Σ is singular in this case. In the limiting case $\delta = 0$, there actually does exist a probability measure μ on $T(4)$ that satisfies (4.2) with the c_{ij} and ξ_i as above. Since the c_{ij} then satisfy (4.14) with the equality sign, that measure μ must be carried by $Z(H) = \{y \in T(4) : y^t H y = 0\}$. In fact, μ assigns mass $1/5$ to each of five points $y^{(r)} \in Z(H)$ ($1 \leq r \leq 5$). Here, $y = y^{(r)}$ has coordinates $y_r = 1/2$; $y_{r\pm 1} = 1/4$; $y_{r\pm 2} = 0$, where the indices are to be interpreted modulo 5.

Recall that $N + S$ is a proper subset of \mathcal{P}_5 since $H \in \mathcal{P}_5/(N + S)$, where H is defined by (4.13). The above construction implies that $N + S$ is not even dense in \mathcal{P}_5 . Namely, consider the 5×5 matrix $C = C(\delta) = (c_{ij})$, defined as in (4.10) and (4.15) with $0 < \delta < \delta_0$ and $c_0 = 1$. The above proof essentially shows that $C(\delta) \in (N + S)^*/\mathcal{P}_5^*$. Hence, \mathcal{P}_5^* is a proper subset of $(N + S)^* = N \cap S$, equivalently, $N + S$ cannot be dense in \mathcal{P}_5 .

REFERENCES

- BERG, C. (1987). The multidimensional moment problem and semigroups. *Proc. Symp. Appl. Math.* **37** American Mathematical Society, Providence, RI. 110–124.
- CASSIER, G. (1984). Problème des moments sur un compact de \mathbb{R}^n et décomposition de polynômes à plusieurs variables. *J. Functional Anal.* **58** 254–266.
- DIANANDA, P. H. (1962). On nonnegative forms in real variables some of all of which are nonnegative. *Proc. Cambridge Philo. Soc.* **58** 17–25.
- HALL, JR., M. (1986). *Combinatorial Theory*. Wiley, New York.
- MARTIN, D. H. AND JACOBSON, D. H. (1981). Copositive matrices and definiteness of quadratic forms subject to homogeneous linear inequality constraints. *Linear Alg. Appl.* **35** 227–258.
- MARTIN, D. H., POWELL, M. J. D. AND JACOBSON, D. H. (1981). On a decomposition of conditionally positive-semidefinite matrices. *Linear Alg. Appl.* **39** 51–59.
- MASERICK, P. H. (1977). Moments of measures on convex bodies. *Pacific J. Math.* **68** 135–152.
- MICCHELLI, A. M. AND PINKUS, A. (1989). Nonnegative polynomials on polyhedra. In *Probability, Statistics and Mathematics, Papers in Honor of Samuel Karlin*. T. W. Anderson, K. B. Athreya and D. L. Iglehart, eds. Academic Press, San Diego, CA. 163–186.

DEPARTMENT OF STATISTICS
RUTGERS UNIVERSITY
NEW BRUNSWICK, NJ 08903

DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF MASSACHUSETTS
AMHERST, MA 01002