GEOMETRY OF REDUCED MOMENT SPACES

S. Karlin and L.S. Shapley

P-97

August 11, 1949

The RAND Corporation
1700 MAIN ST. • SANTA MONICA • CALIFORNIA

GEOMETRY OF REDUCED MOMENT SPACES

S. Karlin and L. S. Shapley

§1. A normalized distribution function on the closed unit interval $[0,1]$ is any real, non-decreasing function $\phi(t)$, continuous to the right, with $\phi(-0) = 0$, $\phi(1) = 1$. The distribution with saltus of one at t and constant elsewhere is denoted by I_t. The <u>n^{th} moment</u> of ϕ is

$$\mu_n(\phi) = \int_{-0}^{1} t^n d\phi(t).$$

The region of n-dimensional euclidean space comprising the points $x_i = \mu_i(\phi)$, $i = 1,2,\ldots,n$, for all distributions ϕ is the <u>reduced moment space</u> $\mathcal{D}^n$. The region comprising the points of the form $x_i = \mu_i(I_t) = t^i$ is a twisted curve, denoted by $\mathcal{C}^n$. $\mathcal{D}^n$ is closed and convex, and for $n \geq 2$ its set of extreme points is precisely $\mathcal{C}^n$.

THEOREM: <u>If x is in the boundary of $\mathcal{D}^n$ then there is a unique distribution ϕ with moments $\mu_i(\phi) = x_i$, and x can be represented as a convex combination of extreme points in just one way, namely</u>

$$x_i = \int_{-0}^{1} \mu_i(I_t) d\phi(t).$$

<u>If x is interior to $\mathcal{D}^n$, then the associated distribution and the convex representation are not unique.</u>

Denote by $\overline{b}(x)$ the number of extreme points participating in the convex representation of a boundary point x of $\mathcal{D}^n$, and $b(x)$ the same but with the points $(0,0,\ldots,0)$ and $(1,1,\ldots,1)$ counted half. Let $L(x)$ be the common intersection of all the supporting planes to $\mathcal{D}^n$ at x; let a(x) denote the dimension of $L(x)$, c(x) the dimension of $L(x) \cap \mathcal{D}^n$.

THEOREM <u>If</u> x <u>is any boundary point of</u> $\mathcal{D}^n$ <u>then</u>

$$2b(x) = a(x) + 1,$$

$$\bar{b}(x) = c(x) + 1.$$

By a suitable extension of the definitions, these relations can be made to hold for points interior and exterior to $\mathcal{D}^n$. In particular, there are two distinct "minimal" representations of the interior points, with $b(x) = (n+1)/2$, corresponding to the representations of two boundary points of $\mathcal{D}^{n+1}$. The boundary of $\mathcal{D}^n$ may be partitioned into 2n disjoint "faces", $\bar{A}^n_a$ and $\underline{A}^n_a$, of surface dimension a, $a = 0,1,\ldots,n-1$, where $a(x) = a$ for $x \in \bar{A}^n_a$ or $\underline{A}^n_a$. c(x) is constant over these faces, and indicates the order of curvature of the surface. For the entire boundary, $a-1 \leq 2c \leq a+1$.

§2. The set of polynomials P(t) of degree n or less which are non negative over $[0,1]$ may be conceived as a conical region in the (n+1)-dimensional euclidean space of the coefficients. If the polynomials are normalized by

$$\int_0^1 P(t)\,dt = 1$$

a convex, closed, n-dimensional region $\mathcal{P}^n$ is obtained. Its extreme points are precisely those polynomials that have n roots (counting multiplicity) in $[0,1]$: they fall into two disjoint components distinguished by the sign of the leading coefficient. The entire region $\mathcal{P}^n$ is spanned by extreme points taken two at a time; in fact:

THEOREM. <u>If</u> P <u>is in</u> $\mathcal{P}^n$ <u>then there is a unique representation</u>

$$P(t) = \alpha \prod_{j=1}^{m} (t-t_{2j-1})^2 + \beta t(1-t) \prod_{j=1}^{m-1} (t-t_{2j})^2, \text{ if } n = 2m,$$

$$P(t) = \alpha t \prod_{j=1}^{m} (t-t_{2j})^2 + \beta(1-t) \prod_{j=1}^{m} (t-t_{2j-1})^2, \text{ if } n = 2m+1,$$

<u>with</u> $\alpha \geq 0$, $\beta \geq 0$, $0 \leq t_1 \leq \dots \leq t_{n-1} \leq 1$. P <u>is interior to</u> $\mathcal{P}^n$ <u>if and only if all of the inequalities are strict</u>.

Denote by $\bar{b}(P)$ the number of distinct roots of $P(t)$; $b(P)$ the same counting roots at 0 and 1 as half. Then for P on the boundary of $\mathcal{P}^n$.

$$\bar{b}(P) = n-a(P)$$
$$2b(P) = n-c(P),$$

$a(P)$ and $c(P)$ being defined geometrically as in §1. The boundary of $\mathcal{P}^n$ may be partitioned into $2n$ disjoint components $\bar{C}^n_c$ and $\underline{C}^n_c$, $c = 0,1,\dots,n-1$, with $c(P) = c$ for P in $\bar{C}^n_c$ or $\underline{C}^n_c$.

§3. A <u>convex cone</u> in $(n+1)$-dimensional space is a set which contains the point $\alpha x + \beta x'$ whenever it contains x and x', α and β being any non negative real numbers [1]. The <u>conjugate</u> cone C^* to a convex cone C is the set of points y which satisfy

$$\sum_{i=0}^{n} x_i y_i \geq 0 \qquad \text{all } x \in C.$$

C^* is convex closed, and if C is closed then $C^{**} = C$.

If $\mathcal{D}^n$ and $\mathcal{P}^n$ are made into convex cones in $(n+1)$-dimensional space by dropping the two normalizing conditions $\int d\phi = 1$ and $\int P dt = 1$, then they are mutually conjugate. Geometrically stated, this means that the coordinates of a boundary point of $\mathcal{D}^n$ are the coefficients of a plane of support to $\mathcal{P}^n$, and vice versa. Interior points correspond to separating planes, exterior points to secant planes. If x is in the boundary of $\mathcal{D}^n$ and if P is interior to the (convex) set of points in the boundary of $\mathcal{P}^n$ which generate planes of support to $\mathcal{D}^n$ at x, then P is said to be <u>conjugate</u> to x.

THEOREM. *If* x *is conjugate to* P, *or* P *conjugate to* x, *then*

$$a(x) + c(P) = c(x) + a(P) = n-1,$$

$$b(x) = b(P), \qquad \bar{b}(x) = \bar{b}(P).$$

The duality of conjugate cones is fundamental in the geometrical interpretation of the theory of two-person, zero-sum games. Motivation for the present study came from the consideration of games described by polynomial payoff functions over the unit square.

§4. In this section, results will be given only for the case $n = 2m$. Analogous results hold for odd dimensions.

The duality just described leads easily to the well-known fact [2], that $\mu_0 = 1, \mu_1, \ldots, \mu_{2m}$ are the moments of a distribution if and only if the quadratic forms

$$\sum_{i,j=0}^{m} x_i x_j \mu_{i+j} \quad \text{and} \quad \sum_{i,j=1}^{m} x_i x_j (\mu_{i+j-1} - \mu_{i+j})$$

are non negative. The first principal minors are

$$\Delta_k = \begin{vmatrix} \mu_0 & \cdots & \mu_k \\ \vdots & & \\ \mu_k & \cdots & \mu_{2k} \end{vmatrix} \quad \text{and} \quad \bar{\Delta}_\ell = \begin{vmatrix} \mu_1 - \mu_2 & \cdots & \mu_k - \mu_{k+1} \\ \vdots & & \\ \mu_k - \mu_{k+1} & & \mu_{2k-1} - \mu_{2k} \end{vmatrix}$$

$k = 0, 1, \ldots, m$, $\ell = 1, 2, \ldots, m$, respectively. These determinants may be used to characterize the faces $\bar{A}_a^{2m}$, $\underline{A}_a^{2m}$ of D^{2m}. In fact, if x is in $\bar{A}_a^{2m}$ then, setting $b = (a+1)/2$,

$$\Delta_k > 0, \quad 1 \le k \le b; \quad \Delta_k = 0, \quad b < k \le m;$$

$$\bar{\Delta}_\ell > 0, \quad 1 \le k < b; \quad \bar{\Delta}_\ell = 0, \quad b \le k \le m.$$

If x is in Δ_a^{2m} then interchange upper and lower bars.

If an interior point $(\mu_1, \ldots, \mu_{2m-1})$ of D^{2m-1} is given, then the extreme values possible for μ_{2m} are found by solving $\underline{\Delta}_{2m} = 0$ and $\bar{\Delta}_{2m} = 0$. The lower value, designated $\underline{\mu}_{2m}$ comes from the first equation, the upper, $\bar{\mu}_{2m}$, from the second. The supporting planes at the opposite boundary points $(\mu_0, \ldots, \underline{\mu}_{2m})$ and $(\mu_0, \ldots, \bar{\mu}_{2m})$ are unique and correspond to the polynomials

$$\left[P_{\underline{\Delta}_m}(t)\right]^2 \quad \text{and} \quad t(1-t)\left[P_{\bar{\Delta}_m}(t)\right]^2 \quad \text{of } P^{2m}, \text{ where}$$

$$P_{\underline{\Delta}_m}(t) = \begin{vmatrix} \mu_0 & \cdots & \mu_{m-1} & 1 \\ \mu_1 & & & t \\ \vdots & & & \vdots \\ \mu_m & \cdots & \mu_{2m-1} & t^m \end{vmatrix}, \qquad P_{\bar{\Delta}_m}(t) = \begin{vmatrix} \mu_1-\mu_2 & \cdots & \mu_{m-1}-\mu_m & 1 \\ \mu_2-\mu_3 & & & t \\ \vdots & & & \vdots \\ \mu_m-\mu_{m+1} & \cdots & \mu_{2m-2}-\mu_{2m-1} & t^{m-1} \end{vmatrix}.$$

The spectra of the unique distributions $\underline{\psi}$ and $\bar{\psi}$ corresponding to these opposite boundary points are the roots of $P_{\underline{\Delta}_m}(t)$ and $t(1-t)P_{\bar{\Delta}_m}(t)$ respectively. The two sets of roots interlock strictly (cf. the representation of (2) and the cumulative weights of the two distributions interlock as well, so that $\underline{\psi}(t) - \bar{\psi}(t)$ has greatest possible number of sign changes.

The polynomials $P_{\underline{\Delta}_k}$, $k = 0,1,\ldots,m$, are orthogonal with respect to the weight factor $d\underline{\psi}(t)$, while $P_{\bar{\Delta}_\ell}$, $\ell = 1,2,\ldots,m$, are orthogonal with respect to $t(1-t)d\bar{\psi}(t)$. These systems of polynomials, together with the analogues for odd n, are the basis for a new geometric approach to the theory of orthogonal polynomials.

THEOREM. <u>There is a 1:1:1 correspondence between (a) the interior of D^{2m-1}, (b) the set of ordered pairs of polynomials $(P_{\underline{\Delta}_m}, P_{\bar{\Delta}_m})$, and (c) the open simplex of strictly interlocking roots</u> $0 = t_0 < t_1 < \ldots < t_{2m} = 1$.

The representation of §2 is thus linked to the polynomials of this section and to the interior of D^{n-1}.

§5. In this section certain geometric features of the reduced moment space D^n and its dual Δ^n will be given. D^n can be inscribed in a very natural simplex S^n whose vertices $x^{(k)}$ are given by

$$x_i^{(k)} = \frac{\binom{k}{i}}{\binom{n}{i}} \qquad i = 1,2,\ldots,n; \; k = 0,1,\ldots,n.$$

Its dual is the simplex inscribed in Δ^n spanned by the polynomials

$$\lambda_{n,k}(t) = \binom{n}{k} t^k (1-t)^{n-k} \qquad k = 0,1,\ldots,n.$$

The barycentric coordinates with respect to S^n of a point x of D^n are

$$\lambda_{n,k} = \int_{-0}^{1} \lambda_{n,k}(t)\,d\phi(t) \qquad k = 0,1,\ldots,n,$$

where ϕ is a distribution having moments $\mu_i(\phi) = x_i$, $i = 1,2,\ldots,n$. This gives a geometric interpretation to the quantities $\lambda_{n,k}$ which play so important a role in the literature on moment theory [3]. The relation of S^n to D^n leads to several qualitative results on moment spaces, and in particular gives a simple geometrical proof of Hausdorf's Theorem.

Another property of D^n is that its center of gravity is given by the continuous distribution $\pi^{-1}(t(1-t))^{-1/2}dt$, whose moments

$$\mu_i = \frac{1 \cdot 3 \cdot \cdots \cdot 2n-1}{2 \cdot 4 \cdot \cdots \cdot 2n}$$

satisfy, for any n, $\mu_n = (\underline{\mu}_n + \bar{\mu}_n)/2$. The supporting planes at $(\mu_1,\mu_2,\ldots,\underline{\mu}_n)$ and $(\mu_1,\mu_2,\ldots,\bar{\mu}_n)$ are parallel, another unique property of this moment sequence. The polynomials $P_{\underline{\Delta}_m}$, $P_{\bar{\Delta}_m}$,

$n = 2m$, for these moments are (up to a constant factor) the Tcheby-cheff polynomials of the first and second kind.

The center of gravity of the simplex S^n is given by the rectangular distribution $\phi(t) = t$, whose moments $1, \frac{1}{2}, \frac{1}{3}, \ldots$ give rise in similar fashion to the Legendre and the first associated Legendre polynomials.

Finally, the n-dimensional volume of S^n has the interesting value

$$V_n = \prod_{i=1}^{n} B(i,i) \ ,$$

where B is the beta function.

6. A similar account can be made for moment spaces based on an arbitrary finite set of moments $\{\mu_i\}$. With certain adaptations, most of the foregoing analysis and geometry can be carried over to infinite moment sequences, and to distributions over the infinite intervals $[0,\infty]$ and $[-\infty,\infty]$. These results and their proof, and additional theorems unstated here, will be published elsewhere.

---

[1] Krein, M. and Šmulian, D., Ann. Math., 41, 556-583 (1940).

[2] Shohat, J. A. and Tamarkin, J. D., The Problem of Moments, American Mathematical Society, New York, 1943, p. 77.

[3] Widder, D. V., The Laplace Transform, Princeton University Press, Princeton, 1941, ch. III.