[redacted]

ON THE GENERAL MOMENT PROBLEM

Richard Bellman and David Blackwell

P-159

7 July 1950

The RAND Corporation
• SANTA MONICA • CALIFORNIA

15

ON THE GENERAL MOMENT PROBLEM

Richard Bellman and David Blackwell

§1. Introduction.

We are interested in the following general problem which we shall first pose abstractly and then illustrate by several concrete examples. Consider a sequence of functions $\{f_n(P)\}$ defined over some space S, and let G be a set of measures defined over a Borel field B of subsets of S.

The moments of g with respect to f_n are defined by

$$a_n = (f_n, g) = \int_S f_n dg . \tag{1}$$

We wish to determine the range of $(a_1, a_2, \cdots, a_n)$ as g varies over G.

Classical examples of this general problem are furnished by the moment spaces defined by

$$a_n = \int_a^b x^n dg(x), \tag{2}$$

or

$$a_n = \int_a^b e^{2\pi i n x} dg(x), \tag{3}$$

where g(x) is a monotone increasing function of bounded variation over the interval [a, b]. Three distinct cases are those where [a, b] is [0, 1], [0, ∞], or [$-\infty$, ∞].

Another well-known moment space is defined by

$$a_n = \int_a^b g(x) f_n(x) dx, \tag{4}$$

where $\|g\| = \left(\int_a^b |g|^p dx\right)^{1/p} \leq 1$, and $\int_a^b |f_n(x)|^{p'} dx < \infty$ for each n, with $p' = p/(p-1)$.

Let M be a Borel measurable set in S. A function defined over S is called the characteristic function of M if it is equal to 1 on M and 0 elsewhere. We shall consider the problem of determining the region of N-dimensional space swept out by the moments

$$a_n = \int_0^1 \phi(x) x^n dx, \tag{5}$$

where $\phi(x)$ varies over the characteristic functions of Borel subsets of the interval $[0, 1]$. That this region is convex and closed is a special case of a result of Liapounoff which tells us that the moment sequence generated by (5) is equivalent to that obtained by letting $\phi(x)$ range over the space of Borel functions assuming only values between 0 and 1.

We shall also treat the trigonometric moment problem, where the moments are given by

$$a_n = \int_0^1 \phi(x) e^{2\pi i n x} dx. \tag{6}$$

First we shall present a method which theoretically resolves the general problem, and then we shall show how it provides the solution to the problem of characterizing the moment space of characteristic functions. The determination of the moment spaces generated by (2), (3) and (4), as well as the classical necessary and sufficient condition due to F. Riesz, [5], that a sequence

a_n be a moment sequence generated by (4), follows readily from our method, but we shall not include these results here, since a treatment of (2) has been given by Karlin and Shapley recently, [4], and (4) can be handled by other methods, cf. Banach, [1].

Our methods, applied to the 2-dimensional analogues of (2) and (5), namely the sequences

$$a_{mn} = \int_0^1 \int x^m y^n dG(x, y), \tag{7}$$

and

$$a_{mn} = \int_0^1 \int \phi(x, y) x^m y^n dx\, dy, \tag{8}$$

where $\phi(x, y)$ is a two-dimensional characteristic function, reduce the characterization problem to problems of algebraic geometry which we shall discuss briefly below.

We should like to add a word concerning the origin of the problem. In connection with our work on games of bluffing, cf. [2] and also an interesting extension given in [6], we encountered the problem of solving the variational problem involved in evaluating

$$\operatorname*{Max}_{\phi} \operatorname*{Min}_{\psi} \int_0^1 \int K(x, y)\phi(x)\psi(y)dx\, dy \tag{9}$$

where K is a known function of x and y and ϕ and ψ are characteristic functions. Due to an optimistic slip, we concluded that K was a polynomial in x and y. In this case, setting

$$a_n = \int_0^1 x^n \phi(x)dx, \qquad b_n = \int_0^1 y^n \psi(y)dy, \tag{10}$$

the problem reduces to that of finding

$$(11) \qquad \operatorname*{Max}_{a_m} \operatorname*{Min}_{b_n} \left[\sum c_{mn} a_m b_n + \sum d_m a_m + \sum e_n b_n \right]$$

where the a_m and b_n range over the moment space defined by (10).

Since there exists an elegant mapping method of solving this latter problem, recently developed by Dresher, [3], the problem of (9), for the case where K is a polynomial in x and y, can be considered as resolved once the moment space has been determined.

In reality, K is a discontinuous function of x and y, and the above shifting of responsibility is not possible. However, the problem of determining the moment space of characteristic functions seemed so intriguing that we persisted nevertheless.

§2. The General Method.

The method we present below is a simple and obvious one and must certainly be contained in essence or entirety somewhere in the literature. Nevertheless, we have not seen it stated in its generality anywhere. There is certainly a close connection between it and the method used by Haar in treating the generalized Chebychev approximation problem. It is applicable whenever, as in the above examples, the moment space is convex, which is always true if the ground space of the g's is convex.

The general problem then is one of determining the boundary of an N-dimensional region R which is known to be convex. Our method is to take a plane – a line in two dimensions – and move

it parallel to itself until a boundary point is reached. This limiting plane will be a supporting plane. Once the supporting planes have been determined, the surface itself is readily obtained using the theory of envelopes. In general, the boundary will not be represented by one analytic formula, but will be pieced together of parts of different analytic surfaces.

Let us without loss of generality consider the case where the functions involved are real and the inner product is real valued. From (1) we obtain

$$(1) \qquad \sum_{n=1}^{N} \lambda_n a_n = \left(\sum_{n=1}^{N} \lambda_n f_n, \; g \right) .$$

We now make the following crucial

Assumption. It is known how to find a $g \in C$ furnishing the maximum of $\left(\sum_{n=1}^{N} \lambda_n f_n, \; g \right)$ for fixed λ_n.

Under the shield of the above assumption, let us set

$$(2) \qquad M(\lambda) = \operatorname*{Max}_{g \in C} \left[\sum_{n=1}^{N} \lambda_n f_n, \; g \right] .$$

It is clear then that the supporting planes are given by

$$(3) \qquad \sum_{n=1}^{N} a_n \lambda_n = M(\lambda) ,$$

where the quantity $M(\lambda)$ we recognize as being proportional to the distance from the origin to the plane given by (3). It is only necessary to obtain maxima, since minima may be obtained by changing

the signs of all the coefficients, λ_n.

In any part of the parameter space in which $M(\lambda)$ is differentiable, we have as the parametric representation of a piece of the boundary

$$(4) \qquad a_k = \frac{\partial M(\lambda)}{\partial \lambda_k}, \qquad k = 1, 2, \cdots, N.$$

Fortunately, we need not worry about the legitimacy of (4) since actually it is the information concerning the properties of the maximizing functions which will enable us to obtain a simple parametric representation of the boundary of the convex region.

The method is illustrated below in consideration of the moment space of characteristic functions.

§3. The Moment Space of Characteristic Functions.

It is easily seen by inspection that the moment space defined by the first two moments

$$(1) \qquad a_0 = \int_0^1 \phi(x)dx$$

$$a_1 = \int_0^1 x\phi(x)dx$$

consists of the region interior to the two curves $a_2 = a_1^2/2$, $a_2 = a_1 - a_1^2/2$, plus the boundary curves. The first case of difficulty is that of three moments, which we shall now treat. Once we see the solution in this case, the general solution follows speedily.

Let us set

$$(2) \qquad a_0 = \int_0^1 \phi(x)dx$$

$$a_1 = \int_0^1 x\phi(x)dx$$

$$a_2 = \int_0^1 x^2\phi(x)dx .$$

We have

$$(3) \qquad \lambda_0 a_0 + \lambda_1 a_1 + \lambda_2 a_2 = \int_0^1 (\lambda_0 + \lambda_1 x + \lambda_2 x^2)\phi(x)dx$$

whence

$$(4) \qquad M(\lambda) = \int_{\lambda_0+\lambda_1 x+\lambda_2 x^2 \geq 0} (\lambda_0 + \lambda_1 x + \lambda_2 x^2)dx,$$

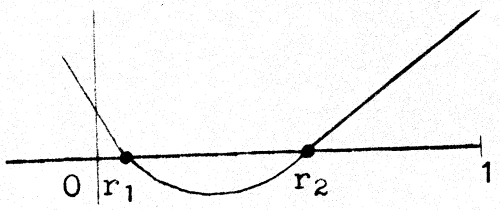
while the extremal functions are defined by

$$(5) \qquad \phi(x) = 1 \quad \text{for} \quad \lambda_0 + \lambda_1 x + \lambda_2 x^2 \geq 0,$$

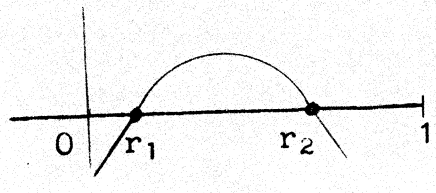
$$= 0 \quad \text{otherwise.}$$

There are consequently two cases, of which all the others are limiting forms, namely

$$(6) \qquad y = \lambda_0 + \lambda_1 x + \lambda_2 x^2$$



$$y = \lambda_0 + \lambda_1 x + \lambda_2 x^2$$



Let us consider the first. Here

$$(7) \qquad \phi(x) = 1 \quad \text{in} \quad [0, r_1] \text{ and } [r_2, 1], \qquad 0 \le r_1 \le r_2 \le 1,$$
$$= 0 \quad \text{otherwise.}$$

In the second,

$$(8) \qquad \phi(x) = 1 \quad \text{in} \quad [r_1, r_2],$$
$$= 0 \quad \text{otherwise.}$$

At this point we realize that in place of using the λ_i as our parameters, we can introduce a great simplification by using the roots r_i as the parameters. In each particular application of the general method, the form of the extremal functions will suggest in a natural way the correct parameters to employ.

We obtain then for one bounding region

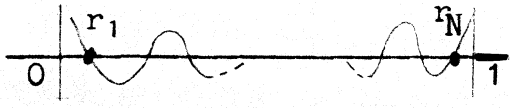
$$(9) \qquad a_0 = \int_0^{r_1} dx + \int_{r_2}^1 dx = 1 + r_1 - r_2,$$
$$a_1 = \int_0^{r_1} x\,dx + \int_{r_2}^1 x\,dx = \frac{1}{2} + \frac{r_1^2 - r_2^2}{2},$$
$$a_2 = \int_0^{r_1} x^2\,dx + \int_{r_2}^1 x^2\,dx = \frac{1}{3} + \frac{r_1^3 - r_2^3}{3}, \qquad 0 \le r_1 \le r_2 \le 1,$$

and for the other

$$(10) \qquad a_0 = r_2 - r_1,$$
$$a_1 = \frac{r_2^2 - r_1^2}{2},$$
$$a_3 = \frac{r_3^3 - r_1^3}{3}.$$

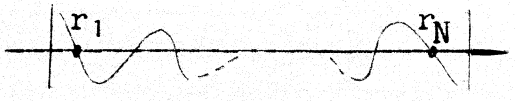
After these preliminary examples, the characterization of the general moment space is simple. It is necessary only to ascertain the manner in which the roots of a general polynomial $P_N(x) = \lambda_0 + \lambda_1 x + \cdots + \lambda_N x^N$ with real coefficients can subdivide the interval $[0, 1]$. This depends upon the degree of the polynomial. If N is even, we have

(17)



together with the result obtained by changing the signs of all the coefficients of the polynomial. If N is odd, we have

(18)



plus again the mirror image in the x-axis.

We can now state our result:

Theorem: If N is even, the boundary of the moment space defined by

$$a_k = \int_0^1 x^k \phi(x)dx, \qquad k = 0, 1, \cdots, N, \tag{19}$$

where ϕ is a characteristic function, consists of two analytic parts, one of which is given by

$$(20)\qquad a_0 = r_1 + (r_3 - r_2) + \cdots + (1 - r_N),$$

$$\vdots$$

$$a_k = \frac{r_1^{k+1}}{k+1} + \left(\frac{r_3^{k+1} - r_2^{k+1}}{k+1}\right) + \cdots + \left(\frac{1 - r_N^{k+1}}{k+1}\right),$$

$$\vdots$$

$$a_N = \frac{r_1^{N+1}}{N+1} + \left(\frac{r_3^{N+1} - r_2^{N+1}}{N+1}\right) + \cdots + \left(\frac{1 - r_N^{N+1}}{N+1}\right),$$

where $0 \le r_1 \le r_2 \le \cdots \le r_N \le 1$, and the other by

$$(21)\qquad a_0 = (r_2 - r_1) + (r_4 - r_3) + \cdots + (r_N - r_{N-1}),$$

$$\vdots$$

$$a_k = \left(\frac{r_2^{k+1} - r_1^{k+1}}{k+1}\right) + \left(\frac{r_4^{k+1} - r_3^{k+1}}{k+1}\right) + \cdots + \left(\frac{r_N^{k+1} - r_{N-1}^{k+1}}{k+1}\right),$$

$$\vdots$$

$$a_N = \left(\frac{r_2^{N+1} - r_1^{N+1}}{N+1}\right) + \left(\frac{r_4^{N+1} - r_3^{N+1}}{N+1}\right) + \cdots + \left(\frac{r_N^{N+1} - r_{N-1}^{N+1}}{N+1}\right).$$

If N is odd, we have for one boundary

$$(22)\qquad a_k = \frac{r_1^{k+1}}{k+1} + \left(\frac{r_3^{k+1} - r_2^{k+1}}{k+1}\right) + \cdots + \left(\frac{r_N^{k+1} - r_{N-1}^{k+1}}{k+1}\right), \quad k = 0,1,\cdots,N,$$

and for the other

$$(23)\qquad a_k = \left(\frac{r_2^{k+1} - r_1^{k+1}}{k+1}\right) + \cdots + \left(\frac{1 - r_N^{k+1}}{k+1}\right), \quad k = 0, 1, \cdots, N.$$

A function yields a boundary point if and only if it has the form

$$(24)\qquad \Phi(x) = 1, \text{ in } [0, r_1], [r_2, r_3], \cdots, [r_N, 1], \quad N \text{ even},$$

$$= 0 \quad \text{elsewhere},$$

or $1 - \Phi$. If N is odd, it must be defined as

(25) $\phi(x) = 1$ in $[0, r_1]$, $[r_2, r_3]$, $\cdots$, $[r_{N-1}, r_N]$,

$= 0$ elsewhere,

or as $1 - \phi$.

§4. The Moment Space of Characteristic Functions of Several Variables.

Although there is no difficulty in following the same procedure when considering the moment space of characteristic functions of several variables, various difficulties arise to prevent an easy solution. Consider as the simplest case, characteristic functions of two variables. We define

$$(1) \qquad a_{mn} = \int_0^1 \int \phi(x, y) x^m y^n dx\, dy$$

and wish to determine the region swept out by the a_{mn}, when m and n independently range over (0, M) and (0, N). As before, we form the expression

$$(2) \qquad \sum \lambda_{mn} a_{mn} = \int_0^1 \int \left(\sum \lambda_{mn} x^m y^n\right) \phi(x, y) dx\, dy.$$

To obtain the maximum we set

$$(3) \qquad \phi = 1 \quad \text{where} \sum_{m,n=0}^{M,N} \lambda_{mn} x^m y^n \geq 0,$$

$= 0$ elsewhere.

It seems then that the problem is reduced to one of algebraic geometry. To obtain the boundary of our moment region, it is necessary to specify the most general subdivision of the unit square by a polynomial curve

$$(4) \qquad \sum_{m,n=0}^{M,N} \lambda_{mn} x^m y^n = 0 .$$

For the first few moments, $m + n \leq 2$, this may be done easily by means of analytical geometry.

Similarly, in connection with the moment space defined by

$$(5) \qquad a_{mn} = \int_0^1 \int x^m y^n dG(x, y) ,$$

we are led to the problem of determining the location of the maxima of a general polynomial

$$(6) \qquad P = \sum_{m,n=0}^{M,N} \lambda_{mn} x^m y^n .$$

In each case, our polynomials have real coefficients.

There are many interesting problems of this type of a mixed algebraic, analytic and geometric nature that arise in conjunction with the investigation of the moment problems arising from different ground spaces, which we propose to consider subsequently.

§5. The Trigonometric Moment Problem.

Let, as in the preceding, $\phi(x)$ be a characteristic function. The method we have applied indicates very clearly that the boundary

of the convex region spanned by the moments

$$(1) \qquad a_n = \int_a^b \phi(x) f_n(x)\,dx\,,$$

depends upon the location of the roots of the general linear combination, $\sum_{n=1}^{N} \lambda_n f_n(x) = 0$, where the λ_n are arbitrary real constants.

The simplest case, treated above, is that where $f_n(x) = x^n$. The next simplest, and essentially equivalent case, is that where $f_n(x) = e^{2\pi i n x}$. In lieu of considering the moments

$$(2) \qquad a_n = \int_0^1 \phi(x) e^{2\pi i n x}\,dx,$$

we consider separately the real and complex parts,

$$(3) \qquad \alpha_n = \int_0^1 \phi(x) \cos 2\pi n x\,dx,$$

$$\beta_n = \int_0^1 \phi(x) \sin 2\pi n x\,dx\,.$$

It follows from previous considerations that in the first case the extremum functions are given by

$$(4) \qquad \phi(x) = 1 \quad \text{for} \quad \sum_{n=0}^{N} \lambda_n \cos 2\pi n x \geq 0\,,$$

$$= 0 \quad \text{elsewhere,}$$

and in the second case by

$$(5) \qquad \phi(x) = 1 \quad \text{for} \quad \sum_{n=1}^{N} \lambda_n \sin 2\pi n x \geq 0,$$

$$= 0 \quad \text{elsewhere.}$$

Since $\cos 2\pi nx = \cos 2\pi n(1-x)$, we can choose as the roots of $\sum_{n=0}^{N} \lambda_n \cos 2\pi nx = 0$, N arbitrary numbers $r_1, r_2, \cdots, r_N$ in the interval $[0, 1]$, and their reflections in $x = \frac{1}{2}$, $1-r_1, 1-r_2, \cdots, 1-r_N$. $\phi(x)$ will then be one in $[0, r_1]$, $[r_2, r_3]$, $\cdots$, etc., or $[r_1, r_2]$, $[r_3, r_4]$, $\cdots$, and zero in the complementary intervals. We have a similar result for (5), and there is no difficulty in writing out the counterpart to our theorem for the ordinary moments.

REFERENCES

1. S. Banach, Théorie des Opérations Linéaires, Warsaw, 1932.

2. R. Bellman and D. Blackwell, Some two-person games involving bluffing, Proc. Nat. Acad. Sci. U.S.A. 35(1949), 600-5.

3. M. Dresher, oral communication. However, his results will appear in print in the near future.

4. S. Karlin and L. Shapley, Geometry of reduced moment spaces, Proc. Nat. Acad. Sci. Vol. 35 (1949), pp. 673-7.

5. F. Riesz, Les systémes d'équations linéaires à une infinité d'inconnues, Paris, 1913.

6. A. Wald, J. Wolfowitz, and A. Dvoretsky, Proc. Nat. Acad. Sci., 1950.