COWLES COMMISSION DISCUSSION PAPER: Mathematics 402 LPC 7(6

Summary of Theory of Polyhedral Cones and Linear Inequalities

by

David Gale, Princeton University

October 6, 1949

Introduction

A number of problems in econometrics are concerned essentially with solutions of systems of linear inequalities. A typical example is the "linear programming" problem in which one wishes to maximize a linear form subject to linear inequalities. Another is the problem of determining the value and optimal strategies for a two-person zero-sum game. Because of their importance in econometrics, therefore, it seems desirable to summarize the fundamental mathematical facts concerning such systems of inequalities. This paper represents an attempt to present this material in as simple and unified a form as possible.

As the title suggests, the theory of inequalities is equivalent to what we have called the theory of polyhedral cones. This theory is nothing more than a geometric interpretation of the inequality theory. However, it has been found that it is often helpful to be able to "visualize" results in some geometric form, and that certain facts which seem obscure from the algebraic point of view become "intuitively obvious" when looked on geometrically. For this reason we have tried whenever possible to state all results twice, first geometrically as theorems about polyhedral cones, and second algebraically as properties of linear inequalities.

The results listed here are given often without proof. In all such cases references are given so that anyone desiring to derive the theorems may refer to the appropriate place in the literature. The principal source for proofs is an article by H. Weyl, "Elementare Theorie der Konvexen Polyeder" [1].

Although no attempt has been made at mathematical completeness as far as proofs are concerned, we have tried to list results in a logical order and particularly to arrange the material in a form that will be easy to remember. The first section of the paper is devoted to definitions needed for the cone theory, in section II the main properties of cones are obtained and interpreted algebraically. The final section involves a few further results of the theory and one or two applications of the theory are given to indicate its use in attacking specific problems.

I. Definitions and Notation

Notation: We shall be dealing with an n dimensional real vector space V which may be thought of as Euclidean n space. The vectors of V will be denoted by small letters $u, v, x, y, \ldots$, and if u and v are two vectors then $u \cdot v$ will denote their inner product,

$$u \cdot v = \sum_{i=1}^{n} u_i v_i \ .$$

The vectors in V will be considered as column vectors, while row vectors will be denoted by primed letters u', v', etc.

Finally we use inequality signs as follows:

$u > v$ means $u_i > v_i$ for all i

$u \geq v$ means $u_i \geqq v_i$ for all i and $u_i > v_i$ for some i

$u \geqq v$ means $u_i \geqq v_i$ for all i .

A. Convex Cones.

A set C contained in V is called a <u>convex cone</u> if: (1) $v_1 + v_2$ are in C whenever v_1 and v_2 are in C; (2) if v is in C and λ is a non-negative real number then λ v is in C.

B. Sum, Intersection, and Polar Cone.

If C_1 and C_2 are convex cones, the sum $C_1 + C_2$ is defined to be the set of all vectors expressible as the sum of vectors from C_1 and C_2.

If C_1 and C_2 are convex cones the intersection $C_1 \cap C_2$ is defined to be the set of all vectors belonging to both C_1 and C_2.

If C is a convex cone, the polar cone C^* is defined to be the set of all vectors u such that $u \cdot v \geqq 0$ for all v in C. We may think of C^* as the set of all vectors making a non-obtuse angle with every vector of C.

One may verify at once that sums, intersections and polars of convex cones are again convex cones.

C. Rays and Halfspaces.

Two special cases of convex cones are of importance. We introduce them now.

A convex cone generated by a single vector v is called a ray or halfline and is denoted by (v) where (v) is the set of all vectors which are non-negative multiples of v.

The polar cone of the ray (v) is called a halfspace and in our notation is denoted by $(v)^*$. It consists of all vectors u such that $u \cdot v \geqq 0$, or geometrically all vectors making a non-obtuse angle with v.

D. Polyhedral Cones.

We are now prepared to introduce the concept of polyhedral cones. We shall give two definitions which look quite different but will later be seen to be equivalent.

(1) C is a polyhedral cone if it is the sum of a finite number of rays, $C = \sum_{i=1}^{k} (v_i)$. C then consists of all non-negative linear combinations of the vectors $v_1, \ldots v_k$.

(2) C is a polyhedral cone if it is the intersection of a finite number of halfspaces, $C = \bigcap_{j=1}^{\ell} (u_j)^*$. C is then all vectors making non-obtuse angles with each of the vectors $u_1, \ldots u_\ell$.

These two definitions may be stated in a more algebraic and less geometric form as follows.

(1') Let A be a k by n matrix. Then C is a polyhedral cone if it consists of all vectors of the form,

$$y'A, \text{ where } y' \geqq 0 .$$

One sees at once the connection between (1) and (1') by taking for the rows of A in (1') the vectors v_i of (1).

(2') Let B be an ℓ by n matrix. Then C is a polyhedral cone if it consists of all vectors x such that

$$Bx \geqq 0 .$$

Letting the rows of B in (2') be the vectors u_j of (2) one immediately sees the equivalence of definitions (2) and (2').

E. Linear space, dimension and type of a polyhedral cone.

The smallest linear space containing a cone C is denoted by [C] and consists of all vectors expressible as the difference of factors in C. The dimension (rank) of C is the dimension of [C] or equivalently the maximum number of linearly independent vectors in C. If C is described by the matrix A of (1') then the dimension of C is the rank of A.

We may also consider the largest linear space contained in C. It consists of all vectors v such that both v and -v lie in C. The dimension of this space is called the type of C. If C is given by the matrix B of (2') then it can be shown that the type of C is n minus the rank of B.

The set of all vectors perpendicular to the polyhedral cone C is denoted by $C^{\perp}$ and consists of all vectors u such that $u \cdot v = 0$ for all

$v \in C$, (for $\in$ read "contained in").

F. Supporting halfspace and hyperplane, extreme halfspace.

If C is a polyhedral cone and $u \in C^*$ then $(u)^*$ is called a supporting halfspace for C. Clearly $(u)^*$ is simply a halfspace containing C. The linear space $(u)^\perp$ is called a supporting hyperplane for C. Now suppose C is given by definition (1), that is $C = \sum_{i=1}^{k} (v_i)$ and let C have dimension n. A supporting halfspace $(u)^*$ is called an <u>extreme halfspace</u> if for n-1 linearly independent vectors from among the $\{v_i\}$ we have $u \cdot v_i = 0$. In other words $(u)^*$ is an extreme halfspace if $(u)^\perp$ contains n-1 independent vectors from among the v_i.

With these definitions in mind we are prepared to discuss the properties of polyhedral cones.

II. Main Properties of Polyhedral Cones

From this point on since all cones discussed will be polyhedral we shall omit the adjective and refer to them simply as cones. Proofs of most of the statements of this section can either be found in Weyl [1] or can be obtained as simple consequences of his theorems. Most of the results follow readily from the fundamental theorem (Hauptsatz) of Weyl's paper which we now state in our own terminology.

<u>WEYL'S THEOREM</u>: If C is an n dimensional cone in n space which is a sum of rays, $C = \sum_{i=1}^{k} (v_i)$ then it is also the intersection of its extreme halfspaces.

We restate the theorem in matrix form as follows: Let the cone C be all vectors of the form

$$x'A \quad \text{where} \quad x' \geqq 0 \quad \text{and} \quad A \quad \text{has rank } n.$$

Then there exists a matrix B such that C is the set of all x for which

$Bx \geqq 0$, and each row of B is orthogonal to n-1 linearly independent rows of A.

If one visualizes Weyl's theorem, say in three space, the result should appear extremely plausible. Nevertheless the proof given by Weyl is non-trivial. Referring to the previous section now we see that Weyl's theorem shows that every n dimensional cone which satisfies the sum definition (1) also satisfies the intersection definition (2). By trivial considerations one can prove this latter statement without the restriction that C be n dimensional. Thus half of the equivalence of definitions (1) and (2) has been established. In order to prove the other half we need the following important fact.

THEOREM 1. If C is a cone and C^* is its polar then $(C^*)^*$ (the polar of C^*) is identical with C.

The proof of this theorem is very simple and we give it here.

Proof. First, C is contained in $(C^*)^*$ for if $x \in C$ then for any $y \in C^*$ we have $x \cdot y \geq 0$, but this means $x \in (C^*)^*$.

Second, $(C^*)^*$ is contained in C for by Weyl's theorem the cone C is an intersection of halfspaces. Therefore if v is not in C there exists a halfspace $(y)^*$ such that C lies in $(y)^*$ and v does not. But since C lies in $(y)^*$ we have $x \cdot y \geq 0$ for all $x \in C$, hence $y \in C^*$, whereas $v \cdot y < 0$ so v is not in $(C^*)^*$.

The above may be thought of as a duality theorem, showing that the relation between a cone and its polar is symmetric.

As an important consequence of the theorem we have

Corollary. Let $a_0, a_1, \ldots a_m$ be m+1 vectors in n-space with the property that whenever $a_i \cdot x \geqq 0$, $i = 1 \ldots m$ then $a_0 \cdot x \geqq 0$. Then a_0 is a non-negative linear combination of the vectors $a_1, \ldots a_m$.

Proof. Let $C = \sum_{i=1}^{m} (a_i)$. The corollary asserts for every

$x \in C^*$, $a_0 \circ x \geqq 0$. This means $a_0 \in (C^*)^*$ and hence $a_0 \in C$ and is therefore expressible as

$$a_0 = \sum_{i=1}^{n} \lambda_i a_i, \qquad \lambda_i \geqq 0.$$

Using theorem 1 we easily show that every "intersection" cone is also a "sum " cone as follows.

Let $C = (u_1)^* \cap \ \ldots \ \cap (u_\ell)^*$. Then v is in C if and only if $u_j \circ v \geq 0$ for $j = 1, \ldots, \ell$, whence $C = \left(\sum_{j=1}^{\ell} (u_j)\right)^*$. But by Weyl's theorem $\sum_j (u_j) = (v_1)^* \cap \ \ldots \ \cap (v_m)^*$ for some set of vectors $v_1, \ldots v_m$. Thus $C = ((v_1)^* \cap \ \ldots \ \cap (v_m)^*)^*$ which is the same as $C = ((\sum_{i=1}^{m} (v_i))^*)^* = \sum_{i=1}^{m} (v_i)$ by theorem 1. We have now shown the complete equivalence of definition (1) and (2).

We next list the important relations between the operations of sum, intersection and polar which now follow easily from preceding results.

Properties

1. The set of all cones is "closed" with respect to the operations $+, \cap, *$. This means that applying any of the operations to cones leads again to a cone. This follows at once by using the appropriate definition (1) or (2) of a cone.

2. $(C_1 + C_2)^* = C_1^* \cap C_2^*$.

This property follows at once from the definitions of the operations involved.

3. $(C_1 \cap C_2)^* = C_1^* + C_2^*$.

4. $(C^*)^* = C$.

Property 4 is simply the statement of theorem 1. Property 3 follows from 2 and 4, for

$$\begin{aligned}(C_1 \cap C_2)^* &= ((C_1^*)^* \cap (C_2^*)^*)^* = ((C_1^* + C_2^*)^*)^* \\ &= C_1^* + C_2^* .\end{aligned}$$

The equivalence of the two definitions of a cone together with properties 1-4 are the fundamental tools used in proving theorems about cones or linear inequalities. The four properties are easily remembered if one thinks of the polar operation as a "mapping" of the set of cones onto itself which interchanges sums and intersections and which when iterated takes each cone back onto itself.

III. Further Properties and Applications.

A. Dimension and Type.

We will show the relation between dimension and type of a cone as defined in section II.

THEOREM 2. $\dim C + \text{type } C^* = n = \text{type } C + \dim C^*$.

Proof. Let y belong to the largest linear space in C^*. Then y and $-y$ lie in C^* so for any $x \in C$, $y \cdot x \geqq 0$ and $(-y) \cdot x = -(y \cdot x) \geqq 0$, hence $y \cdot x = 0$ so y is in $C^{\perp}$. On the other hand if $y \in C^{\perp}$ then y and $-y$ are in C^* so the largest linear space contained in C^* is exactly $C^{\perp}$. Hence if $\dim C = p$, then $\dim C^{\perp} = n - p = \text{type } C$ proving the first equality of the theorem, and the second follows by Theorem 1. (From this we get at once the expression in terms of rank for the type of a cone given in I,E).

B. Non-negative Cones.

As another application we prove the following fact about cones. By the <u>positive octant</u> P we mean all vectors $v \geqq 0$. The <u>negative octant</u> is denoted by $-P$.

Theorem 3. If the cone C intersects the negative octant $-P$ only in zero then C^* contains a vector v interior to P (i.e. $v > 0$).

Proof. Let Q be the cone $P + C$. Then clearly Q is not the whole space V since it does not intersect $-P$, and therefore Q^* is not empty (use Theorem 1). By property 2 of cones $Q^* = P^* \cap C^*$. Since Q^*

is a cone it is the sum of rays $Q^* = (v_1) + \dots + (v_k)$. Let $v = v_1 + \dots + v_m$. Then v lies in Q^* hence in C^*. Now let p lie in P. Then $p \cdot v \geqq 0$, and if $p \cdot v = 0$ then $p \cdot v_i = 0$ for $i = 1 \dots m$ hence p and $-p$ would lie in $(C^*)^{\perp}$ which is in $(C^*)^* = C$. But by assumption $-p$ lies in C only if $p = 0$. It follows that for $p \geq 0$, $v \cdot p > 0$ from which we see that $v > 0$.

The above theorem can be expressed in matrix form as follows:

If A is any matrix then either there exists $x' \geq 0$ such that $x'A \leq 0$ or there exists $x > 0$ such that $Ax \gneqq 0$.

This is called the "theorem of the alternative for a matrix" in "Theory of Games and Economic Behavior" and is used in proving the fundamental theorem of the two-person game.

C. Existence of Value of a Symmetric Game.

The main theorem of the two-person game proves the existence of a value and optimal mixed strategies for the two players. It is not difficult to reduce this to proving the theorem for the case when the game is symmetric, that is, when the pay-off matrix is skew-symmetric. Stated algebraically the theorem we desire to prove is the following.

THEOREM 4. If A is a skew-symmetric matrix then there exists a vector x' such that $x'A \geqq 0$, $x' \geq 0$, $\sum_{i=1}^{n} x'_i = 1$.

This theorem is an easy consequence of theorem 3. However in order to illustrate the use of the cone theory we prefer to derive it directly from the main properties of cones.

Proof. Note that it suffices to find x' satisfying (1) $x'A \geqq 0$ and (2) $x' \geq 0$ since we can then satisfy $\sum_i x'_i = 1$ by multiplying x' by the appropriate constant. Now suppose no x' exists satisfying (1) and (2). Let C be the cone consisting of all vectors of the form $x'A$ with $x' \geq 0$. If P is the positive octant then we have assumed $C \cap P = 0$. Now we assert $(-C) + P$ is not all of V (where $(-C)$ denotes all vectors whose negatives lie in C).

In particular if $q \geqq 0$ then $-q$ is not in $(-C) + P$ for then we would have $-q = -y + p$ with $y \in C$, $p \in P$, or $y = p + q \in P$ contradicting $C \cap P = 0$. Since $(-C) + P$ is not all of V it follows that $((-C) + P)^* \neq 0$, for if it were then we would have $((-C) + P)^{**} = (0)^* = V = (-C) + P$. But $((-C) + P)^* = (-C)^* \cap P^* = (-C)^* \cap P$ for from the definition of polar one sees that $P^* = P$. This means that there exists a vector $x \geq 0$ such that $-x \in C^*$ which says $A(-x) = -Ax \geqq 0$ or $Ax \leqq 0$. But since A is skew-symmetric $Ax = -x'A$ so $-x'A \leqq 0$, or $x'A \geqq 0$ (and $x' \geq 0$) contrary to the assumption that no such x' existed. This proves the theorem.

D. Application to Linear Programming.

A useful result on cones, which however does not seem to follow from our previous results is the following (essentially theorem 2 of Weyl).

THEOREM 5. If $C = \sum_{i=1}^{k} (v_i)$ is a cone and v lies in C then it is possible to write v as a positive linear combination of not more than n vectors from among the v_i.

This theorem has application to problems in linear programming. In particular it is used by Dantzig [3] to show that if "feasible solutions" exist then one can be found depending on not more than n points.

Stated in matrix form theorem 5 reads. If $v = x'A$ where $x' \geq 0$ and A is a k by n matrix, $k > n$, then there exists x'' such that $v = x''A$, $x'' \geq 0$ and at least $k - n$ of the components of x'' are zero.