* *This research was supported by National Science Foundation Grant Number GU-2059.*

GENERATION OF THE ADMISSIBLE BOUNDARY OF A CONVEX POLYTOPE

by

Richard Shachtman
Department of Statistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 743

February 24, 1971

Generation of the Admissible Boundary of a Convex Polytope*

Richard Shachtman
University of North Carolina at Chapel Hill
Chapel Hill, North Carolina 27514

Abstract

The purpose of this paper is to develop a constructive algorithm for generation of the admissible boundary of a convex polytope, when the polytope is the convex hull of a given (finite) set of points as well as to mention some related results. One application is to the sensitivity of decisions to changes in the prior distribution for problems in statistical decision theory.

1. Introduction

Problems in linear optimization are often reduced to examining the boundary, or part of the boundary, of a convex polyhedral set, P. In particular, problems in the related areas of linear programming, game theory, and statistical decision theory require this sort of search. In some cases, the equations (or half-spaces) defining P may be given and the objective is to determine one or more "optimal" vertices (extreme points) of P. In other models, a set of points A may be given and the convex hull of A, C(A), is of interest. If C(A) is compact and has a finite number of extreme points, then C(A) is a bounded (convex) polyhedral set, or, as we will refer to such a set here, a polytope; see Grünbaum [5, chp. 3]. The results here are directed predominantly towards this second class of problems.

The main purpose of this paper is to develop a constructive algorithm for generation of the so-called admissible boundary, $\mathcal{A}$(A), of a polytope

* *This research was supported by National Science Foundation Grant Number GU-2059.*

$P = C(A)$, as well as to mention some related results. An immediate application follows if one is interested in maximizing a non-negative linear functional f on P, since it then suffices to consider $A(A)$; see Section 2. Of course, if f is fixed, one searches for the point $a \in A$ for which $f(a)$ is a maximum. However, in many contexts f is an approximation to a "true" linear functional, or is formulated from subjective estimates. This is the case, for example, in applications of statistical decision theory, or to be more precise, when one must assess a prior distribution for certain unknown states of nature in a decision analysis[1]; see Section 5.

Thus, constructing the admissible boundary should be useful in analyzing problems involving non-negative linear functionals. Here, these models are reduced to linear programs and the efficiency and flexibility of the simplex algorithm and of existing post-optimality techniques is brought to bear. It is thought that the admissible boundary approach should be useful in other contexts where objective functions, or transformations of them, have non-negative coefficients or, in general, in studying the structure of convex polytopes; for example, an application may be made to integer programming problems such as those discussed in Elmaghraby [3].

Background and notation is covered in Section 2. In Sections 3 and 4, an algorithm is indicated which is a constructive method for generating a certain convex set containing A, called the Bayes hull of A, which has the same admissible boundary as the polytope $C(A)$. The description of $A(A)$ is obtained from that of the Bayes hull of A. This turns out to be more efficient than working directly with $C(A)$. Moreover, for an "optimal" point, or set of points, one may identify sets of adjacent extreme points. The application to decision analysis is described in Section 5, as well as an interpretation for the

---

1 *Background and further references for decision analysis may be found in Howard [8, 9, 10] and Raiffa [13].*

associated linear program and its dual. Results similar to those in Section 2 are known (e.g., see Blackwell and Girshick [2]); brief new proofs are supplied to make the presentation self-contained and easily accessible.

The algorithm is being programmed; initial efficiences are indicated, but extended comments on computational experience must be reported later.

2. Background and Notation

Definition 2.1. We will denote the convex hull of a set A in real m-space, $\mathbb{R}^m$, by $C(A)$ and the transpose of $x \in \mathbb{R}^m$ by x'.

Suppose $S = \{v_1, \ldots, v_n\}$ is a set of points in m-space; each v_j is a (column) vector $(v_{1j}, \ldots, v_{mj})'$, $j = 1, \ldots, m$. A linear functional f defined on $\mathbb{R}^m$ is maximized over S if and only if it is maximized over the polytope $C(S)$. Moreover, we may and do assume, without loss of generality, that S is contained in the upper orthant of $\mathbb{R}^m$, $\{x \in \mathbb{R}^m: x \geq 0\}$.

Definition 2.2. A linear functional $f: \mathbb{R}^m \to \mathbb{R}$ is non-negative if $f(x) \geq 0$ for all $x \geq 0$, $x \in \mathbb{R}^m$. We will write $f \geq 0$ and all functionals will be linear. Equivalently, f is non-negative on $\mathbb{R}^m$ if and only if $f(x) = a'x = \sum_1^m a_i x_i$ and $a_i \geq 0$ for all $i = 1, \ldots, m$.

Definition 2.3. Let x and $y \in \mathbb{R}^m$. Then x dominates y if $x_i \geq y_i$, $i = 1, \ldots, m$ and the inequality is strict for some i_0. For $A \subset \mathbb{R}^m$, $x \in A$ is an admissible point of A if it is not dominated by any other point in A. If A is finite, we say that such an x is pure admissible. If A is convex, we say that such an x is mixed admissible, or, simply, admissible. Denote the set of admissible points of A by $M(A)$.

Definition 2.4. A point $y \in C(S)$ is said to be on the admissible boundary of $C(S)$ if it is contained in a supporting hyperplane of $C(S)$, $H = \{x \in \mathbb{R}^m: a'x = b\}$, such that $b > 0$ and $a > 0$ (i.e., $a_i > 0$ for $i = 1, \ldots, m$). We write $A(S) = A(C(S))$ for the admissible boundary of $C(S)$. This simplified structure will suffice for our use. For a slightly different setting and alternative proof of the following, see Blackwell and Girshick [2].

Corollary 2.5. $A(S) = M(S)$.

Proof. $y \in A(S) \Rightarrow$ there exists $H = \{x: a'x=b\}$ such that $a > 0$, $b > 0$, $a'y = b$ and $a'x \leq b$ for all $x \in C(S)$. If $y \notin M(S)$, there exists $z \in C(S)$ such that $z \geq y$ and $z_i > y_i$ for some i. Then $b \geq a'z > a'y$. Hence $A(S) \subset M(S)$. Conversely, $y \notin A(S) \Rightarrow$ for all H with $a > 0, b > 0$ and $a'y = b$, there exists $z \in C(S)$ with $a'z > b$. Write $y = (y_1,\ldots,y_m)'$ and for each $y_r > 0$, say $r \in R \neq \phi$, define a by $a_i = 0$ for $i \neq r$, $a_r = y_r$, and $b = a'y = y_r^2$. This generates (at most m) hyperplanes $H_r = \{x: a'x=b\}$, $r \in R$, and points $z^{(\ell)} = (z_{1\ell},\ldots,z_{m\ell})' \in C(S)$ with $a'z^{(\ell)} > y_r^2$ or $z_{r\ell} > y_r$. Let $z = \sum \alpha_\ell z^{(\ell)}$, where $0 < \alpha_\ell < 1$; then $z \in C(S)$ and $z = \sum \alpha_\ell (z_{1\ell},\ldots,z_{m\ell})' \geq y$ where the inequality is strict for $r \in R$. Hence $y \notin M(S)$. □

Theorem 2.6. If $f \geq 0$ on $\mathbb{R}^m$, then $\max\{f(x): x \in A(S)\} = \max\{f(x): x \in C(S)\}$.

Proof. Let d and c be the maximum values f attains on $A(S)$ and $C(S)$, respectively. (The values are attained since $C(S)$ is compact.) Hence, $c = f(y)$ for some $y \in C(S)$. Define $H = \{x: f(x)=c\}$ and note that it is a supporting hyperplane of $C(S)$. We will show that H contains a point w (not necessarily equal to y) with $w \in A(S)$. Therefore since, obviously, $d \leq c$, $f(w) = c = d$.

If $f(x) = a'x$ and $a > 0$, we are finished. Otherwise, let $K \subset \{1,\ldots,m\}$ such that $a_i = 0$ for $i \in K$ and $a_i > 0$ for $i \notin K$. Let $F = H \cap C(S)$. Define $w = (w_1,\ldots,w_m)'$ by the following:

(1) For each successive index $i \in K$, set the linear program $w_i = \max x_i$, subject to $x \in F$ and $w_j = \max x_j$, $j < i$.

(2) For $i \notin K$, $w_i = y_i$ where $y = (y_1,\ldots,y_m)'$.

Suppose there exists $z \in C(S)$ such that $z \geq w$ and $z_j > w_j$ for some j. $f(z) = \sum_{i \notin K} a_i z_i \geq \sum_{i \notin K} a_i w_i = f(w) = c \geq f(z) \Rightarrow z \in H \Rightarrow z \in F$ and $j \in K$. Consequently, $w_j = \max x_j \geq z_j$. Contradiction. Hence w is admissible. □

3. The Bayes Hull

Definition 3.1. $\mathcal{X}_i = \{x \in \mathbb{R}^m: x_i = 0\}$ is the i^{th} coordinate hyperplane, $i = 1,\ldots,m$.

We use the following matrix for the points in S:

(i,j)	v_1	...	v_n
	v_{11}	...	v_{1n}
	⋮		
	v_{m1}	...	v_{mn}

where we have retained only those points in S which are pure admissible. Also, we make the translation insuring that $S \subset \{x \in \mathbb{R}^m: x \geq 0\}$ and that at least one entry in each row of the matrix is zero; i.e., there is at least one point of S in each $\mathcal{X}_i$.

Definition 3.2. Compute $u_i = \max\{v_{ij}: j = 1,\ldots,n\}$, $i = 1,\ldots,m$ and define the "corner" points $c_i = (0,\ldots,0,u_i,0,\ldots,0)'$, $i = 1,\ldots,m$. $C = \{c_i: i = 1,\ldots,m\}$. Assume $u_i > 0$, for all i.

Definition 3.3. Fix i, $i = 1,\ldots,m$, and observe that $u_i = v_{ik}$ for some $k = k(i)$, say $v_k = (v_{1k},\ldots,v_{mk})'$. Define P_i to be the set of (point) projections of v_k on each coordinate hyperplane $\mathcal{X}_r$, $r \neq i$. That is $P_i = \{p_{rk}: r \neq i\}$ and $p_{rk} = (p_1,\ldots,p_m)$ with $p_i = v_{ik}$ for $i \neq r$ and $p_r = 0$. $P = \bigcup_1^m P_i$.

Note that P may include some points in S and/or C.

Definition 3.4. Let the polytope $\mathcal{B}(S) = C(S \cup C \cup P \cup \{0\})$ be called the Bayes (convex) hull of S, or $\mathcal{B}$-hull.

<u>Corollary 3.5.</u> Construct the (unique) hyperplane H_0 through the corner points $C = \{c_i\}$. Then $H_0 \cap B(S) = C(C)$.

<u>Theorem 3.6.</u> $A(B(S)) = A(C(S)) = A(S)$.

<u>Proof.</u> $y \in A(S) \Rightarrow y \in C(S)$ and there exists $H = \{x: a'x=b\}$, $a > 0$, $b > 0$ such that $a'y = b$ and $a'x \le b$ for all $x \in C(S)$.

If $x \in B(S)$, then $x = \sum\alpha_j z_j$, $\alpha_j \ge 0$, $\sum\alpha_j = 1$ and $z_j \in S \cup C \cup P$. For $z_j \in C$, $z_j = (0,\ldots,0,u_i,0,\ldots,0)'$, for some i, where $u_i = v_{ik}$ and $k = k(j)$; hence, $0 \le z_j \le v_k$, $v_k \in S$. For $z_j \in P$, $z_j \in P_i$, for some i and $z_j = p_{r\ell}$, for some $r \ne i$, $\ell = \ell(j)$; hence, $0 \le z_j \le v_\ell \in S$. Therefore, $x = \sum\alpha_j z_j \le \sum_S \alpha_j z_j + \sum_C \alpha_j v_{k(j)} + \sum_P \alpha_j v_{\ell(j)} = \sum\alpha_j v_{h(j)}$, $v_{h(j)} \in S$, and $a'x \le \sum\alpha_j a'v_{h(j)} \le b$.

Thus $y \in B(S)$ and H is such that $a'y = b$ and $a'x \le b$ for all $x \in B(S)$. Consequently, $y \in A(B(S))$; i.e., $A(S) \subset A(B(S))$.

$y \in A(B(S)) \Longrightarrow y \in B(S)$ and H exists as above. But $a'x \le b$ for $x \in B(S)$ implies the same for $x \in C(S)$. Hence, it suffices to show $y \in C(S)$. $y = \sum\alpha_j z_j$, $\alpha_j \ge 0$, $\sum\alpha_j = 1$ and $z_j \in S \cup C \cup P$. If $y \notin C(S)$, there exists j such that $z_j = c_\ell$ [or $p_{r\ell}$] for some $\ell = \ell(j)$ [or r and ℓ]. In either case, we see from the first part of the proof that $z_j \le v_k \in S$, for some $k = k(j)$, with the inequality strict for at least one component. Thus $b = a'y = \sum\alpha_j a'z_j < \sum\alpha_j a'v_{k(j)} \le b$. Contradiction. As a result, $y \in A(S)$; i.e., $A(B(S)) \subset A(S)$. □

The last, combined with Theorem 2.6, tells us that we may construct $B(S)$ whenever we wish to maximize an f over S and perform any sensitivity analysis. Moreover, we will see that this construction will, in general, reduce computational effort in generating $A(S)$ by use of the algorithm in the next section. In particular, surfaces of the polytope $C(S)$ which are not admissible are not generated; some $\{v_j\}$ which are pure admissible but not (mixed) admissible (including some extreme points of $C(S)$) do not appear in the iterations of the algorithm and the number of constructions is reduced.

The following guarantees that using the hyperplane $H_0 = \{x: a_0'x=b_0\}$ through the corner points $C = \{c_i\}$ for the initial construction in the algorithm does not exclude any points in the admissible boundary.

Theorem 3.7. $A(S) \subset \{x: a_0'x \geq b_0\}$.

Proof. Let $y \in A(S)$ and suppose that $a_0'y < b_0$. Then $y \in C(C\cup\{0\})$ and $y \notin C$. $y = \sum\alpha_j c_j$, $\alpha_j \geq 0$, $\sum\alpha_j = 1$ and for some j, say $j = \ell$, $0 < \alpha_\ell < 1$. As before, $c_j \leq v_k$, for some $k = k(j)$, and $y \leq \sum\alpha_j v_{k(j)} = z$ with the component $y_\ell = \alpha_\ell u_\ell = \alpha_\ell v_{\ell k(\ell)} < \sum\alpha_j v_{\ell k(j)}$, or y is dominated by the point $z \in C(S)$. Contradiction. □

Corollary 3.8. All extreme points of $B(S)$, except 0, are in $\{x: a_0'x \geq b_0\}$.

4. The Algorithm

The procedure which defines an algorithm for the construction of the polytope $B(S)$ from the set S consists of starting with the hyperplane H_0, searching for the extreme point (or points) of $B(S)$ which is (are) "most distant" from H_0 and generating hyperplanes intersecting $H_0 \cap B(S) = C(C)$ and the point(s). The procedure is repeated until, at the final iteration, the hyperplanes defining $B(S)$ have been generated. This collection includes those hyperplanes defining the admissible boundary, $A(S)$.

Definition 4.1. Let F be the intersection of a polytope P with a supporting hyperplane. If a polytope F has dimension $k \in \{0,\ldots,m\}$, it is a k-face. 0-faces are vertices; 1-faces are edges; $(m-1)$-faces are called facets and P is determined by a finite number of (hyperplanes defining) facets; see Grünbaum [5, chp. 3].

Let T denote $S \cup C \cup P$, with points $z_j \in T$, j in some (finite) set J.

<u>Theorem 4.2.</u> Let $H = \{x: a'x = b\}$ be a hyperplane through m points in T and suppose there exists at least one point $z \in T$ such that $a'z > b$. Write $\beta = max\{a'z_j: z_j \in T\}$. Then either

(1) there exists exactly one point, z_k, satisfying $a'z = \beta$ and z_k is an extreme point of $\mathcal{B}(S)$; or

(2) there is more than one such point, say $E = \{z_k: k \in K\}$ is the set of points with $a'z_k = \beta$, and $C(E)$ is a face of $\mathcal{B}(S)$.

In (2), if $E \subset S$, the face $C(E)$ is in $A(S)$.

<u>Proof.</u> (1) If z_k is the unique point satisfying $a'z = \beta$, then $a'z_j < a'z_k$ for all $j \neq k$. Suppose z_k is not an extreme point, then $z_k = \sum\alpha_r z_r$, $\alpha_r \geq 0$, $\sum\alpha_r = 1$, $\{z_r: r \in R\}$ a subset of the extreme points of $\mathcal{B}(S)$, and for some r, say $r = s$, $0 < \alpha_s < 1$. Hence, $a'z_k = \sum\alpha_r a'z_r < \sum\alpha_r a'z_k = a'z_k$. Contradiction.

(2) $\beta = a'z_k > a'z_j$ for all $k \in K$, $j \notin K$. The points in E either determine a hyperplane, H_1, which is parallel to H, or lie in H_1. In either case, $H_1 = \{x: a'x = \beta\}$. Observe that for all $k \in K$, $a'z_k \geq a'z_r$ for all extreme points $z_r \in \mathcal{B}(S)$, $r \in R_1$. Thus for all $y \in \mathcal{B}(S)$, $y = \sum\alpha_r z_r$ and $a'y = \sum\alpha_r a'z_r \leq a'z_k = \beta$. That is, E lies in a supporting hyperplane of $\mathcal{B}(S)$, as does $C(E)$. □

After establishing $T = S \cup C \cup P$, the procedure is to start with the hyperplane H_0 and determine a set $E = \{z_k: k \in K\}$ as in the preceeding theorem. (3.7 and 3.8 guarantee that this excludes no nontrivial extreme points of $\mathcal{B}(S)$ and, in particular, no admissible points.) If K is a singleton, $E = \{z_0\}$, construct the m new hyperplanes determined by each collection of m-1 corner points and z_0. If K is not a singleton, then form all new hyperplanes determined by each collection of m-1 corner points and each point in E. For each hyperplane at this stage, and for subsequent stages, repeat the above procedure. Each stage involves the construction of at most $\binom{m+1}{m}-1 = m$ new hyperplanes.

Theorem 4.3. The algorithm converges. At any stage of the algorithm, the procedure locates and includes at least one new extreme point, or face, of $\mathcal{B}(S)$. The set of extreme points is finite. The algorithm terminates when, at some stage, for all hyperplanes $H = \{x: a'x = b\}$ there exist no points $z \in T$ with $a'z > b$. These hyperplanes, in addition to the coordinate hyperplanes $\{X_i\}$, include those hyperplanes which determine all facets of $\mathcal{B}(S)$.

In a computer program for the procedure, one stores (the equation of) a hyperplane containing $C(E)$ if the cardinality of K is m or more. Note that such a listing may include hyperplanes determining k-faces with $k < m - 1$. This is not necessarily a problem since the equations just provide redundant information; e.g., this doesn't affect post-optimality analysis. However, removing these may decrease computational requirements and is useful in identifying extreme points. If there are not too many such cases, the following lemma suggests a technique for identifying the facets.

Lemma 4.4. A hyperplane $H = \{x: a'x = b > 0\}$ is determined by $\{x_j\}_1^m$ if and only if the $\{x_j\}_1^m$ are linearly independent.

Proof. The result is known for arbitrary b when the set in the sufficiency is replaced by $\{x_j - x_m\}_1^{m-1}$, renumbering if necessary (see Hadley [7]). Hence, the sufficiency with $\{x_j\}_1^m$ follows since $\{x_j\}$ linearly independent $\Rightarrow \{x_j - x_m\}$ linearly independent. For the necessity, let $\sum_1^m \alpha_j x_j = 0$, $x_j = (x_{1j},\ldots,x_{mj})'$ and $\{\alpha_j\}_1^m$ scalars. Let X be the matrix with columns $(x_1,\ldots,x_m)$. Then $X\alpha = 0$. Suppose that $\alpha \neq 0$. Then $det(X) = 0$ and there exists a $\beta = (\beta_1,\ldots,\beta_m)' \neq 0$ with $\beta' X = 0$. Thus, x_j satisfies $\beta' x_j = 0$

for all j; that is, $\{x_j\}^m$ is contained in a hyperplane through the origin. Contradiction. □

It is evident that any $H = \{x: a'x = b\}$ with which we are concerned will have $b > 0$.

An alternative technique for removing redundant constraints may be derived from Rubin [15].

Beyond "mathematical" convergence, the algorithm is thought to be computationally efficient since at each stage, a surface point of the polytope is located and, as mentioned in Section 3, constructions involving pure admissible extreme points which are not admissible are eliminated. This last is achieved by using $\mathcal{B}(S)$ instead of $C(S)$.

If we denote the set of admissible extreme points of $C(S)$ by E, then the total number of stages, or steps, of the algorithm is bounded by the cardinality of $E \cup P$. At each step, the total number of constructions of hyperplanes is at most m. Of course, in general, the number of iterations will be related to the number of facets of $\mathcal{B}(S)$. The so-called upper bound conjecture for the number of facets of a polytope, given the number of vertices, has recently been shown to be true by P. McMullen; see the announcement in Grünbaum [6, p.1183] and formulae in [5], [6] and [11].

Assume S is non-degenerate in $\mathbb{R}^m$ (not contained in a hyperplane).

Lemma 4.5. v is an extreme point of $C(S)$ if and only if there exist (at least) m facets of $C(S)$ containing v.

Proof. Write $F_k = H_k \cap C(S)$, $k = 1,\ldots,K$, for the (distinct) facets which determine $C(S)$. Suppose $v \in \bigcap_{k=1}^r F_k$ and $v \notin \bigcup_{k>r} F_k$, $r < m$ (renumbering, if necessary). Let $\gamma = \frac{1}{2}\min\{||v-H_k||: k > r\} > 0$, $0 < \varepsilon < \gamma$ and $N_\varepsilon(v) = \{x\in\bigcap_{k=1}^r F_k: ||x-v|| < \varepsilon\}$. The dimension of $\bigcap_{k=1}^r H_k = m-r \geq 1$; hence, one can show that points in $N_\varepsilon(v)$ may be selected which have m-r components free to range in m-r intervals, thus contradicting that v is an extreme point.

Suppose $0 \neq v \in F = \bigcap_1^m F_k = \bigcap_1^m H_k \cap C(S)$. Each F_k has dimension m-1. If $dim(F_j \cap F_k) = m-1$, $j \neq k$, then $dim(H_j \cap H_k) = m-1 \iff$ the dimension of the solution space of the equations defining H_j, H_k is m-1 $\Rightarrow$ the equations are redundant $\Rightarrow H_j = H_k \Rightarrow F_j = F_k$. But the facets are distinct; hence, $dim(F_j \cap F_k) \leq m-2$. By induction, $dim(\bigcap_1^r F_k) \leq m-r$, $r \leq m$, and $dim(F) = 0$. 0-faces are vertices and $F = \{v\}$. □

<u>Corollary 4.6.</u> (i) The final form of the constraint set defining $B(S)$ may be written as the matrix equation $Ax \leq 1$ where $A = \binom{A_1}{A_2}$, $1 = (1,\ldots,1)'$, the rows of the submatrix A_1 are the normals to hyperplanes defining the facets of $A(S)$ and the rows of the submatrix A_2 are the normals to hyperplanes defining the facets of $B(S)$ which contain at least one (extreme) point in $C \cup P$.

(ii) If we maximize $f \geq 0$ subject to $Ax \leq 1$, the linear program is consistent, bounded and the solution is in $A(S)$.

(iii) We may list the admissible extreme points of $C(S)$ $[B(S)]$ corresponding to each facet defined in A_1 [A].

When the algorithm has terminated, all hyperplanes will be of the form $H = \{x: a'x = b > 0\}$; those with $a > 0$ correspond to facets of $A(S)$ and those with at least one $a_i = 0$ correspond to non-admissible facets of $B(S)$.

For each hyperplane, there is a corresponding set of points of S. Many are already identified as extreme points (whenever an E-set is a singleton or doubleton); the others may be tested and identified by using 4.5 (whenever a point satisfies m or more hyperplanes defining facets). Hence, for an "optimal" point, or set of points, one may identify sets of "adjacent" extreme points; these sets are useful for post-optimality analysis.

Techniques discussed in Balinski [1] may be used to locate all extreme points of a polytope when the set of defining equations, $Ax = b$, is given. In our model, $Ax = b$ is to be generated and much of the effort in identifying extreme points is expended in the generation; hence, it is thought that little advantage is to be gained by using the techniques of [1]. However, computational comparisons depend on the particular polytope of interest. In examples for which the E-sets are always singletons or doubletons, all extreme points are quickly located by our procedure on the first "pass" without recourse to further checking.

5. Applications

As mentioned, one application of the admissible boundary approach is to decision analysis. In the sequel, we rely on a description of the normal form of analysis for statistical decision theory; see, for example, Raiffa [13, chp. 6] and Raiffa and Schlaifer [14, chp. 1]. For applications to decision analysis, see Howard [8][2], [9].

We assume that part of a decision analysis model is described by a decision tree, namely the probabilistic (nature's) tree of the normal form. (The dual mode of analysis is by the extensive form via the chronological tree; both modes lead to the same final decision.) Each path through the tree describes a possible strategy, s_j, $j = 1,\ldots,n$, and has been assigned a certain utility value given a particular state of nature, θ_i, $i = 1,\ldots,m$; i.e., we write $v_{ij} = v(s_j|\theta_i)$ to denote the utility value for strategy s_j when θ_i is the "true" state of nature.

To this point we have assumed only that the above values have been assigned and, consequently, any analysis is independent of prior distribution assessments. Hence, a major advantage of normal form analysis is that prior distribution input occurs at the final stage of the analysis thus reserving subjective judgements until the end, and dealing only with admissible strategies (as few strategies as possible).

---

2 *In [8, p.39] reverse the labels for Figures 4 and 5; Figure 4 is nature's tree and Figure 5 is the chronological decision tree.*

The vector $v_j = (v_{1j},\ldots,v_{mj})'$ represents m possible values for selecting strategy s_j, the vector of joint conditional values. If $S = \{v_j: j = 1,\ldots,n\}$ represents the set of values for all strategies, then $C(S)$ is the set of all mixed strategies and if we accept the basic axioms of preference (utility) theory, our objective is to maximize the expected utility $EU = \sum_{i=1}^{m} p_i u_i = p'u$ over the polytope $C(S)$, where $p = (p_1,\ldots,p_m)'$ is a prior distribution. As we have seen, it suffices to maximize EU over $A(S)$.

It is often the case that the decision maker is uncertain about the precise values for his prior distribution on the unknown states of nature, or (perhaps to a lesser extent) about the utility values $\{v_{ij}\}$. If he establishes tentative estimates for p and the $\{v_{ij}\}$, we may solve the above linear program and use post-optimality analysis to determine classes of probability distributions for which the decision remains the same. (The $\{v_{ij}\}$ may be varied by modifications on the $A_1 v = 1$ describing $A(S)$ in Cor. 4.6.) Alternative approaches, where values are assumed to be fixed, are suggested in Fishburn, Murphy and Isaacs [4] and Pierce and Folks [12]. In their main approach, the decision maker determines a "nearest" probability distribution for which his decision changes and uses this as a guideline for the sensitivity of his decision to his initial estimates for p. Comparative efficiencies of the procedures appear to be problem-dependent.

The linear program of 4.6 and its dual may be interpreted for decision analysis. Let (P) be $\max\ p'x$, s.t. $A_1 x \le 1$, $x \ge 0$ and (D) be $\min\ y'\cdot 1$, s.t. $y'A_1 \ge p'$, $y \ge 0$. p and x have dimension $m\times 1$; y is $r\times 1$ and A is $r\times m$ where m is the number of states of nature and r is the number of facets which determine $A(S)$. Each such facet F is associated with a (unique) set, $\sigma(F)$, of joint conditional values corresponding to certain (pure) strategies in S. (We identify s_k and $v_k = (v_{1k},\ldots,v_{mk})'$.)

For any point $x = (x_1,\ldots,x_m)' \in A(S)$, x_i is the amount of utility value given the state of nature θ_i. x must be in at least one facet F of $A(S)$, thus $x = \sum_{k\in K}\alpha_k v_k$, where $\alpha_k \geq 0$, $\sum\alpha_k = 1$ and $v_k \in \sigma(F)$. Hence x_i is the amount of value deriving from these strategies with proportions α_k for each strategy s_k, given the state of nature θ_i.

If $x^* = (x_1^*,\ldots,x_m^*)'$ is an optimal basic solution to (P), then $p'x^* = y^{*\prime}1 = \sum_{j=1}^{r} y_j^*$, where $y^* = (y_1^*,\ldots,y_n^*)'$ is an optimal solution to (D). By complementary slackness, if the j^{th} constraint in (P) is non-binding (the corresponding slack variable is positive), then $y_j^* = 0$; thus the only y_j^* which may be in the solution at a positive level are those associated with facets F_j whose hyperplanes are binding. (If some F_j is binding and the corresponding $y_j^* = 0$, this just means that all the alternate optima in (P) are in other facets F_k with corresponding $y_k^* > 0$.) Consequently, an interpretation for the dual variables y_j^* is that $y_j^*[\sum_{k=1}^{r} y_k^*]^{-1}$ is the proportion of value derived from $\sigma(F_j)$, the strategies defining the facet F_j. This interpretation holds for $y_j^* > 0$ or $y_j^* = 0$. Note that a normalized form of the constraint set in (P), e.g., $A_1 x \leq 1$, must be used for this interpretation to be reasonable. Parametric programming may be used to generate sets of proportions of value derived from strategies given various prior distributions.

Modifications on the algorithm will handle the addition or change of strategies.

The algorithm is being programmed and further investigation has the objective of resolving such conjectures as: the number of cases in which the algorithm must resort to identifying facets by independence checks is small in relation to the number of facets of $A(S)$.

Acknowledgment

The author is grateful to Professor Ronald A. Howard whose Decision Analysis Seminar prompted this study.

References

1. M. L. Balinski: "An Algorithm for Finding All Vertices of Convex Polyhedral Sets", *J. Soc. Ind. Appl. Math.*, 9 (1961) 72-88.

2. D. Blackwell and M. A. Girshick: *Theory of Games and Statistical Decisions*, Wiley, 1954.

3. S. E. Elmaghraby: "On Integer Linear Programming by the Reduction Method", *N.C.S. Univ. Opns. Res. Report No. 70* (to appear).

4. P. C. Fishburn, A. H. Murphy and H. H. Isaacs: "Sensitivity of Decisions to Probability Estimation Errors: A Reexamination", *Opns. Res.*, 16 (1968) 254-267.

5. B. Grünbaum: *Convex Polytopes*, Wiley Interscience, 1967.

6. _____: "Polytopes, Graphs, and Complexes", *Bull. Am. Math. Soc.*, 76 (1970) 1131-1201.

7. G. Hadley: *Linear Algebra*, Addison-Wesley, 1961.

8. R. A. Howard: "Bayesian Decision Models for System Engineering", *IEEE Trans. Sys. Sci. Cyb.*, SSC-1 (1965) 36-40.

9. _____: "The Foundations of Decision Analysis", *IEEE Trans. Sys. Sci. Cyb.*, SSC-4 (1968) 211-219.

10. _____: "Decision Analysis: Perspectives on Inference, Decision and Experimentation", *Proc. IEEE*, 58 (1970) 632-643.

11. P. McMullen: "On a Problem of Klee Concerning Convex Polytopes", *Israel J. Math.*, 8 (1970) 1-4.

12. D. A. Pierce and J. L. Folks: "Sensitivity of Bayes Procedures to the Prior Distribution", *Opns. Res.*, 17 (1969) 344-350.

13. H. Raiffa: *Decision Analysis: Introductory Lectures on Choices under Uncertainty*, Addison-Wesley, 1968.

14. H. Raiffa and R. Schlaifer: *Applied Statistical Decision Theory*, M.I.T. Press edition, 1968.

15. D. S. Rubin: "Redundant Constraints and Extraneous Variables in Integer Programs", *Management Sci.* (to appear).