

FACIAL DISJUNCTIVE PROGRAMS AND SEQUENCE OF CUTTING-PLANES

Charles E. BLAIR

University of Illinois. Dept. Business Administration, Urbana, IL 61801, USA

Received 25 May 1979

Revised 18 February 1980

A cutting-plane procedure for facial disjunctive programs is presented. At each step the point to be cut away and the disjunction used to generate the cut may be chosen freely. Some general comments on finiteness proofs for discrete programs are also given.

Introduction

Many discrete optimization problems can be viewed as systems of linear inequalities together with restrictions of an "either-or" type, e.g., either $x_1 = 0$ or $x_5 = 0$ or $x_7 = 0$. Balas [1, 2] introduced facial disjunctive programs to develop the general theory of such problems. $P = \{x \mid Ax \geq b\} \subset \mathbb{R}^n$ is a polytope given by the usual inequality constraints. A facial disjunctive constraint is a requirement that the feasible set satisfy at least one of the inequalities $d_i x \geq e_i$, $i = 1, \dots, k$; where $P \cap d_i x \geq e_i$ is a face of P . The constraints of a facial disjunctive program consist of the inequalities defining P , together with t facial disjunctive constraints, each of the form

$$x \in \bigcup_{i \in D_j} (P \cap d_i x \geq e_i) \quad j = 1, \dots, t. \quad (1)$$

Disjunctive programs include as special cases zero-one integer programs and linear complementarity problems.

We consider cutting-plane methods of obtaining the feasible set S . For $Q \subset P$ the inequality $\alpha x \geq \beta$ is said to be valid for Q if $\alpha z \geq \beta$ for all $z \in Q$. For $1 \leq j \leq t$ define

$$E(j, Q) = \bigcup_{i \in D_j} (Q \cap d_i x \geq e_i) \quad (2)$$

Hence

$$S = \bigcap_{j=1}^t E(j, P) \quad (3)$$

No method of obtaining all valid inequalities for S directly is known. However, Balas [1, see also 3] has shown how valid inequalities for $E(j, Q)$ may be obtained

by solving certain linear programs. It is also shown in [1] that

$$S = E(t, \text{conv}(E(t-1, \text{conv}(E(t-2, \dots E(2, \text{conv}(E(1, P) \dots))).$$

In principle, the feasible set may be obtained by adding all the inequalities generated by the first disjunctive constraint, then adding all cuts generated by the second constraint (applied to $E(1, P)$), and so forth. Since the number of facets of $E(1, P)$ is typically exponential, other methods are needed.

Jeroslow [5] considered schemes in which cutting-planes are added one at a time. One started with $Q_0 = P$. At the k th step one has $Q_k \subset Q_{k-1}$ and an extreme point z_k of Q_k . If $z_k \in S$ the algorithm stops. Otherwise j is determined such that $z_k \notin E(j, Q_k)$. An inequality $\alpha x \geq \beta$ is obtained (using the linear program techniques mentioned above) which is valid for $E(j, Q_k)$ and such that $\alpha z_k < \beta$, i.e., the point z_k is cut away. Then $Q_{k+1} = Q_k \cap \alpha x \geq \beta$, an extreme point z_{k+1} of Q_{k+1} is located, and so forth.

Jeroslow showed that if j, α, β are suitably chosen at each step, then $\text{conv}(S)$ will be obtained in finitely many steps, regardless of the choice of extreme point z_k at each step. The finiteness proof is non-trivial. We give a small example at the end of this paper to show that finiteness may fail if one simply chooses at each step an arbitrary facet of $E(j, Q_n)$ which cuts away z_n .

The problem is posed in [5] whether one can still obtain finite convergence if one is allowed to choose z_k and also the j_k such that $z_k \notin E(j_k, Q_k)$ at each step (i.e., one chooses both the extreme point and the disjunctive constraint to be used to cut it away arbitrarily). We will show that this can be done, although the cuts used may be difficult to compute. Then we present the example illustrating a way in which finiteness may fail. Finally we discuss a general approach to the proof of finiteness for cutting-plane methods on discrete optimization problems.

Preliminary analysis

We begin with a formal description of cutting-plane procedures in general. Let W be the set of all finite sequences of quadruples $(z_i, \alpha_i, \beta_i, j_i)$

$$W = \{ \langle (z_i, \alpha_i, \beta_i, j_i) \rangle_{0 \leq i \leq K} \mid [z_i, \alpha_i \in \mathbb{R}^n; \beta_i \in \mathbb{R}; 1 \leq j_i \leq t] \} \quad (4)$$

such that (i) $\alpha_0 = 0, \beta_0 = \vec{0}$; (ii) z_m is an extreme point of

$$Q_m = P \bigcap_{i=0}^m \alpha_i x \geq \beta_i; \quad (5)$$

(iii) $z_m \notin E(j_m, Q_m)$; (iv) $\alpha_m z_{m-1} < \beta_m$, and (v) $Q_m \supset S$ for all $0 \leq m \leq K$. We will denote those $w \in W$ of length $k+1$ [i.e., last term is $(z_k, \alpha_k, \beta_k, j_k)$] by W^k . Thus, $W = \bigcup_k W^k$.

We identify a cutting-plane procedure with a function $A: W \rightarrow \mathbb{R}^{n+1}$ that assigns to each $w \in W^k$ $A(w) = (\alpha_{k+1}, \beta_{k+1})$ such that (iv) and (v) are satisfied for $m = k+1$. A is finitely convergent if and only if

$$\text{there is no infinite sequence } \langle (z_i, \alpha_i, \beta_i, j_i) \rangle \quad (6)$$

such that, for every m , $w_m \in W^m$ and $(\alpha_m, \beta_m) = A(w_{m-1})$ [w_m = first $m + 1$ members of the infinite sequence].

In other words, regardless of the choice of z_i and j_i at each step, one eventually reaches a situation in which every extreme point of Q_m is in $\bigcap_{j=1}^i E(j, P) = S$, hence $Q_m = \text{conv}(S)$.

A crucial role in our subsequent analysis is played by the fact that $E(j, P)$ is a union of faces of P . Let

$$P(m) = \{x \mid x \in F \text{ for some face } F \text{ of } P \text{ of dimension } \leq m\}. \quad (7)$$

Suppose we have obtained a polytope $Q \supset S$. Since the extreme points of $\text{conv}(S)$ are extreme points of P , S is contained in the convex hull of those extreme points of P which are members of Q , i.e., $\text{conv}(Q \cap P(0)) \supset S$. More generally,

Lemma. Let $Q \supset S$ and $0 \leq m \leq n$. Then $\text{conv}(Q \cap P(m)) \supset S$.

Proof. Since $P(m+1) \supset P(m)$ it suffices to show this for $m = 0$. By (2), (3), and de Morgan's law, we may write S as a union of intersections of faces of P . Since the intersection of faces is a face, this establishes that S is a union of faces of P . Since a face of P is the convex hull of certain extreme points of P , it follows that $\text{conv}(S)$ is the convex hull of extreme points of P , as claimed above. \square

We describe informally our cutting-plane scheme. At each step we have $Q_m \supset S$ and z_m an extreme point of Q_m . Let

$$d_m = 0 \text{ if } z_m \text{ is an extreme point of } P \text{ otherwise } d_m = \text{dimension of the unique face } F_m \text{ of } P \text{ such that } z_m \in \text{interior}(F_m). \quad (8)$$

If $d_m = 0$, we cut z_m away using any inequality valid for S . Since P has a finite number of extreme points this only happens finitely often. If $d_m > 0$, then

$$z_m \notin \text{conv}(Q_m \cap P(d_m - 1)), \quad (9)$$

$$\text{conv}(F_m \cap Q_m \cap P(d_m - 1)) \supset F_m \cap E(j_m, Q_m). \quad (10)$$

(9) follows from the fact that z_m is an extreme point of Q_m . (10) holds because $F_m \cap E(j_m, Q_m)$ is a union of proper faces of F_m .

We construct an inequality $\alpha_{m+1}x \geq \beta_{m+1}$ which is valid for $E(j_m, Q_m)$ and cuts away z_m . To ensure finite convergence we arrange that $F_m \cap \alpha_{m+1}x = \beta_{m+1}$ is (roughly speaking) a facet of $\text{conv}(F_m \cap Q_m \cap P(d_m - 1))$.

The convergence theorem

For $w \in W^k$ and $1 \leq d \leq n$ let

$$L(d, w) = \text{largest } m \leq k - 1 \text{ such that } d_m < d. \quad (11)$$

For $Q \subset P$ and F a d -dimensional face of P a finite set $S(Q, F) \subset R^{n+1}$ is

defined to be a sharp set of inequalities if

$$\text{conv}(Q \cap F \cap P(d-1)) = Q \cap F \bigcap_{(\alpha, \beta) \in S(Q, F)} \alpha x \geq \beta. \quad (12)$$

Sharp sets of inequalities exist, e.g., the facets of $Q \cap P(d-1)$.

Theorem. Suppose that for every F , $Q \cap S(Q, F)$ is a sharp set of inequalities. Suppose $A: W \rightarrow \mathbb{R}^{n+1}$ is a cutting-plane procedure such that for every $w \in W^k$ if $A(w) = (\alpha_{k+1}, \beta_{k+1})$, then

$$\alpha_{k+1} z_k < \beta_{k+1}; \quad (13)$$

$$\text{if } d_k = 0, \quad Q_k \cap \alpha_{k+1} x \geq \beta_{k+1} \supset S; \quad (14)$$

$$\text{if } d_k > 0, \quad \text{then for some } (\alpha, \beta) \in S(Q_{L(d_k, w)+1}, F_k) \quad (15)$$

$$F_k \cap Q_k \not\subset F_k \cap Q_k \cap \alpha x \geq \beta \supset F_k \cap Q_k \cap \alpha_{k+1} x \geq \beta_{k+1};$$

$$\text{if } d_k > 0, \quad Q_k \cap \alpha_{k+1} x \geq \beta_{k+1} \supset Q_k \cap P(d_k - 1). \quad (16)$$

Then A is a finitely convergent procedure, i.e., (6) holds.

Proof. With each $w \in W^k$ we associate $C(w) = (a_0, a_1, \dots, a_n)$, an $(n+1)$ -tuple of natural numbers measuring the complexity of Q_k . a_0 is the number of extreme points of P in Q_k . For $1 \leq d \leq n$ we define

$$a_d = \sum_F N(F) \quad (17)$$

where $N(F)$ is the number of $(\alpha, \beta) \in S(Q_{L(d, w)+1}, F)$ such that $Q_k \cap F \cap \alpha x \geq \beta \not\subset Q_k \cap F$, and the sum is over all d -dimensional faces F of P . Let $w^* \in W^{k+1}$ be such that $(\alpha_{k+1}, \beta_{k+1}) = A(w)$ and w = the first $(k+1)$ terms of w^* . Let $C(w^*) = (a_0^*, \dots, a_n^*)$. If $d_k = 0$ $a_0^* < a_0$ because $z_k \in Q_k - Q_{k+1}$. If $d_k > 0$ and $i \leq d_k$, then $L(i, w) = L(i, w^*)$. Since $Q_{k+1} \subset Q_k$, $a_i^* \leq a_i$. Further $a_{d_k}^* < a_{d_k}$ because, by (15), $N^*(F_{k-1}) < N(F_{k-1})$. Hence $C(w^*)$ is lexicographically smaller than $C(w)$. By well ordering, no infinite sequence is possible. \square

Condition (16) is not used directly in the convergence proof. Its purpose is to insure that, at each step of the algorithm, if $d_k > 0$, then there is some $(\alpha, \beta) \in S(Q_{L(d_k, w)+1}, F_k)$ such that $\alpha z_k < \beta$. This follows from the fact that, for all $d > 0$, $Q_k \supset Q_{L(d, w)+1} \cap P(d-1)$.

Solution of the problem of Jeroslow [5]

We must construct a finitely convergent A such that, for $w \in W^k$ $A(w) = (\alpha_{k+1}, \beta_{k+1})$ is such that

$$Q_k \cap \alpha_{k+1} x \geq \beta_{k+1} \supset E(j_k, Q_k) \quad (18)$$

We require a variant of the separating hyperplane theorem, which can be proved by the usual convex analysis methods.

Lemma. Let $T \subset P$ be a polytope, F a face of P , $v \in \mathbb{R}^n$, $\delta \in \mathbb{R}$, $z \in F - T$. If $vz < \delta$ and $F \cap vx \geq \delta \supset F \cap T$, then there are (α, β) such that $P \cap \alpha x \geq \beta \supset T$ and $(F \cap \alpha x = \beta) = F \cap vx = \delta$.

Now we specify the desired A . $S(Q, F)$ consists of all the facets of $Q \cap P(d-1)$. If $d_k = 0$, we cut z_k away using any valid inequality for $E(j_k, Q_k)$. If $d_k > 0$, then by (9) and (10) there are $(v, \delta) \in S(Q_{L(d_k, F_k)+1}, W)$ such that $vz_k < \delta$ and $Q_k \cap F_k \cap vx \geq \delta \supset \text{conv}(Q_k \cap F_k \cap P(d_k - 1)) \supset Q_k \cap F_k \cap E(j, Q_k)$. We let $T = \text{conv}((Q_k \cap P(d_k - 1)) \cup E(j, Q_k))$ and apply the lemma to obtain $\alpha_{k+1}, \beta_{k+1}$ satisfying (13), (15), (16), and (18).

The cutting-plane procedure described above is impractical because, among other things, each step depends on locating a facet of $\text{conv}(Q_k \cap P(d_k - 1))$. However, the convergence theorem yields finiteness results for many different algorithms, depending on different choices of the sharp sets $S(Q, F)$. Larger sharp sets make each individual cut (α_i, β_i) easier to compute, but the cuts become shallower.

An example of non-convergence.

The finiteness proofs here and in [5] are surprisingly messy. We offer an example of a non-convergent cutting-plane procedure, which suggests that some delicacy is required to insure finiteness.

Let $P \subset \mathbb{R}^3$ be the polytope whose extreme points are

$$\begin{array}{ll} (3, 3, 0) & (3, 3, 5) \\ (0, 8, 0) & (0, 8, \theta) \\ (8, 0, 0) & (8, 0, \phi) \\ (3, 10, 0) & (3, 10, \phi) \\ (10, 3, 0) & (10, 3, \phi) \\ (8, 8, 0) & \end{array} \quad [\theta \text{ and } \phi \text{ to be specified later}]. \quad (19)$$

Geometrically P has a hexagon base and an upper surface that is a "creased hexagon". The two are joined at the point $(8, 8, 0)$.

There are two disjunctive constraints

$$E(1, P) = \{(x, y, z) \mid x = 0 \text{ or } y = 0 \text{ or } z = 0\}, \quad (20)$$

$$E(2, P) = \{(x, y, z) \mid x + y = 6 \text{ or } x = 10 \text{ or } y = 10 \text{ or } z = 0\}.$$

$$S = P \cap E(1, P) \cap E(2, P) = \{x, y, z \in P \mid z = 0\}.$$

Let θ, ϕ satisfy

$$4 < \theta < 5, \quad \phi = \frac{3}{2}\theta. \quad (21)$$

Then the extreme point $(0, 8, \theta)$ can be cut away by the inequality

$$(5 - \phi)x + (5 - \phi)y + 7z \leq 65 - 6\phi. \quad (22)$$

(22) is a facet of $E(2, P)$ which goes through $(3, 3, 5)$; $(3, 10, \phi)$; and $(10, 3, \phi)$.

$Q_1 = P \cap (22)$ has the same extreme points as $Q_0 = P$ except that $(0, 8, \theta)$ and $(8, 0, \theta)$ are replaced by $(0, 8, \theta')$; $(8, 0, \theta')$ such that

$$\phi > \frac{3}{8}\theta', \quad \theta' = \frac{2}{7}\phi + \frac{25}{7}. \quad (23)$$

The extreme point $(3, 10, \phi)$ can be cut away by the inequality

$$\theta'x + \theta'y + 8z \leq 16\theta'. \quad (24)$$

(24) is a facet of $P(1, Q_1)$ which goes through $(0, 8, \theta')$; $(8, 0, \theta')$ and $(8, 8, 0)$.

$Q_2 = Q_1 \cap (24)$ has the extreme points $(10, 3, \phi)$ $(3, 10, \phi)$ replaced by $(10, 3, \phi')$ and $(3, 10, \phi')$ where

$$4 < \theta' < 5, \quad \phi' = \frac{3}{8}\theta'. \quad (25)$$

Since (25) is the same as (21) the process can be continued indefinitely.

Two remarks should be made about this example. At each step there is only one j such that the present extreme point is not in $E(j, Q_k)$. This is not a case of choosing the wrong disjunction but rather the wrong facet, which keeps creating undesirable new extreme points. Secondly, it should be noted that the sequence of extreme points does not approach a member of S as a limit.

Concluding remarks

The finiteness questions are still rather mysterious. Both the methods described here and in [5] have the irritating property that the finiteness proof may fail if deeper cuts than ones specified are used (this is related to the creation of unwanted extreme points in our example). The author feels that present finiteness proofs are more cumbersome than they should be, and that a theory unifying the various techniques is needed.

Many finiteness proofs (perhaps all) are based on the idea that the polytope after a cut is simpler than the polytope before the cut. Different cutting-plane algorithms correspond to different definitions of "simpler". The definition used in this paper was determined by the numbers a_d in (17). The Gomory method of integer forms (in its usual presentation) uses a measure of complexity $(\alpha_0, \dots, \alpha_k)$ where α_0 = largest integer value obtainable for the objective function z by a point in the polytope; α_i = largest integer value obtainable for x_i by a point in the polytope such that $z = \alpha_0$ and $x_j = \alpha_j$ for all $j < i$. We hope to study the various

algorithms and their corresponding definitions of "simpler" (equivalently, complexity measures defined on polytopes) later.

Finally, we wish to mention a question related to Gomory's method of integer forms. Gomory [4], after showing that certain row selection rules guaranteed finite convergence, observed that he knew of no example of non-convergence arising from an arbitrary selection of rows at each step. Twenty years later, no such example has been constructed.

References

- [1] E. Balas, Disjunctive programming: properties of the convex hull of the feasible points, MSRR No. 348, GSIA, Carnegie-Mellon University (1974).
- [2] E. Balas, Disjunctive programming, MSRR No. 415. Discrete Optimization II, Ann. Discr. Math. 5 (North-Holland, Amsterdam, 1979) 3-51.
- [3] C. Blair and R. Jeroslow, A converse for disjunctive constraints, J. Optimization Theory 25 (1978) 195-206.
- [4] R. Gomory, An algorithm for integer solutions to linear programming, IBM Technical Report, 1958.
- [5] R. Jeroslow, A cutting-plane game and its applications, CORE discussion paper no. 7724 (1977).