Faculty Working Papers

DISJUNCTIVE PROGRAMS AND SEQUENCES OF
CUTTING-PLANES

Charles E. Blair, Assistant Professor,
Department of Business Administration

#576

College of Commerce and Business Administration
University of Illinois at Urbana-Champaign

Sharpe, William (1964). "Capital Asset Prices: A Th

Equilibrium Under Conditions of Risk," Journal of F

Stigler, G. (1963). Capital and the Rate of Return in

Industries, Princeton: Princeton University Press.

Taylor, G.R. (1951). The Transportation Revolution 181

York: Holt, Rinehart and Winston.

U.S. Bureau of the Census, (1975). Historical Statistic

Washington, D.C.: Government Printing Office.

FACULTY WORKING PAPERS

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

June 6, 1979

DISJUNCTIVE PROGRAMS AND SEQUENCES OF CUTTING-PLANES

Charles E. Blair, Assistant Professor,
Department of Business Administration

#576

Summary:

We continue work of Balas and Jeroslow on cutting-planes algorithms for disjunctive programs. We show that, in theory, finite convergence can still be obtained even if the extreme point to be cut away and the disjunction used to generate the cut are chosen arbitrarily. A more computational variant of the same idea is also presented as well as a small example illustrating non-convergence.

Disjunctive Programs and Sequences of Cutting-Planes

by

Charles E. Blair

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 disjunctive programs to develop the general theory of such problems. $P = \{x | Ax \geq b\} \subset R^n$ is a polytope given by the usual inequality constraints. A disjunctive constraint is a requirement that the feasible set satisfy at least one of the inequalities $d_i x \geq e_i$, $i = 1,\ldots k$; where $P \cap d_i x \geq e_i$ is a face of P. The constraints of a disjunctive program consist of the inequalities defining P, together with t disjunctive constraints, each of the form

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

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

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

Hence

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

No method of obtaining 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, E(t-1, E(t-2,...E(2, E(1,P))...). 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 kth 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 inquality $\alpha x \geq \beta$ 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 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+1} 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. We then present a related algorithm in which the cuts are generated via linear programs. Finally we present the example mentioned previously and discuss finiteness proofs in general.

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 $(z_i, \alpha_i, \beta_i, j_i)$

(4) $W = \{<(z_i, \alpha_i, \beta_i, j_i)>_{0 \le i \le K}\}$ $[z_i, \alpha_i \in R^n; \beta_i \in R; 1 \le j_i \le t]$

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

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

(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 $0 \le m \le K$. 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 $A: W \to R^{n+1}$ 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

(6) there is no infinite sequence $<(z_i, \alpha_i, \beta_i, j_i)>$

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

(7) $P(m) = \{x | x \in F$ for some face F of P of dimension $\leq m\}$

Suppose we have obtained a polytope $Q \supset S$. Since the extreme points of 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,

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

<u>Proof</u>: 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 conv(S) is the convex hull of extreme points of P, as claimed above. Q.E.D.

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

(8) d_m = dimension of that face F_m of P such that

$$z_m \in \text{interior } (F_m)$$

If $d_m = 0$, i.e., z_m is an extreme point of P, 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

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

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

(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 Conv $(F_m \cap Q_m \cap P(d_m - 1)$.

<u>The Convergence Theorem</u>

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

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

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

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

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

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

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

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

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

$$F_k \cap Q_k \nsubseteq F_k \cap Q_k \cap \alpha x \geq \beta \supset F_k \cap Q_k \cap \alpha_{k+1} x \geq \beta_{k+1}.$$

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

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

Proof: With each $w \in W^k$ we associate $C(w) = (a_0, a_1, \ldots a_n)$, 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

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

where N(F) is the number of $(\alpha,\beta) \in S(Q_{L(d,w)+1}, F)$ such that $Q_k \cap F \cap \alpha x \geq \beta \nsupseteq Q_k \cap F$, 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 $C(w^*) = (a_0^*, \ldots a_n^*)$. 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. Q.E.D.

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

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

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

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

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 $(\nu, \delta) \in S(Q_{L(d_k, F_k)+1}, W)$ such that $\nu z_k < \delta$ and $Q_k \cap F_k \cap \nu x \geq \delta \supset \mathrm{conv}(Q_k \cap F_k \cap P(d_k-1)) \supset Q_k \cap F_k \cap E(j, Q_k)$. We let $T = \mathrm{conv}((Q_k \cap P(d_k-1)) \cup E(j, Q_k))$ and apply the lemma to obtain α_{k+1}, β_{k+1} satisfying (13), (15), (16), and (18).

An Algorithm Based on Cuts Generated by Linear Programs

The cutting-plane procedure described in the preceeding section seems impractical because, among other things, each step depends on locating a facet of $\mathrm{conv}(Q_k \cap P(d_k-1))$. In this section we describe an algorithm based on similar ideas in which the cuts are generated at each step by finding basic feasible solutions (not necessarily optimal) to certain linear programs.

We convert the representation of P to equation form by adding slack variables, so $P = \{y | B_o y = b_o, \ y \geq 0\}$ where B_o has q rows and r columns. At step k of the algorithm the next cut is added by introducing a new row to the constraints and a new non-negative variable $v^{(k)}$. In general we have

(19) $\quad Q_k = \{y | B_k y + C_k v = b_k; \ y, \ v \geq 0\}$ where

B_k is $(q+k)\times r$ and C_k is $(q+k)\times k$. $B_k^{(i)}$ and $C_k^{(i)}$ will denote the ith columns of the matrices, while $y_k^{(i)}$, $v_k^{(i)}$ will be components of vectors.

At each step $(y_k,v_k) \in R^{r+k}$ is an extreme point of Q_k corresponding to a choice of basic variables in (19). y_k is in the interior of the face of P

$$(20) \qquad F_k = \{y \mid B_o y = b_o,\ y \geq 0\ y^{(i)} = 0 \text{ if } y_k^{(i)} = 0\}$$

As before d_k = dimension of F_k and $L(d,k)$ = the largest $m \leq k-1$ such that $d_m < d$.

At step k+1 (y_k,v_k) must be cut away. Let

$$(21) \qquad J = \{i \mid y_k^{(i)} > 0\}$$

$$(22) \qquad D_k = \bigcup_{i \in J} \{(y,v) \mid (y,v) \in Q_k \text{ and } y^{(i)} = 0\}$$

Then $D_k \cap F_k = F_k \cap P(d_k-1) \cap Q_k$ and $\text{conv}(D_k) \supset \text{conv}(Q_k \cap P(d_k-1)$. It follows from results in [3] that every inequality $\alpha y + \beta v \geq \nu$ valid for $Q_m \cap D_k$ $(0 \leq m \leq k)$ can be obtained from a feasible solution to the "disjunctive dual program"

$$(23) \qquad \text{minimize } \alpha y_k + \beta v_k - \nu$$

subject to

$$\alpha_j \geq u_i B_k^{(j)} \quad j = 1,\ldots r;\ i \in J - \{j\}$$

$$\beta_j \geq u_i C_k^{(j)} \quad j = 1,\ldots m \quad i \in J$$

$$\beta_j = u_i C_k^{(j)} \quad m < j \leq k \quad i \in J$$

$$\nu \leq u_i b_k \quad i \in J$$

$$(-1, \ldots, -1) \leq u_i \leq (1, \ldots, 1)$$

In (23) the variables are $\alpha \in R^r$, $\beta \in R^k$, $\nu \in R$, and $u_i \in R^{q+k}$ for all $i \in J$. The inequalities corresponding to basic feasible solutions to (23) constitute a sharp set of inequalities $S(Q_m, F_k)$.

If $(y_k, v_k) \notin S$ we cut it away by using an (α, β, ν) corresponding to a basic feasible solution of (23), $m = L(d_k,k)+1$, with negative objective function value (it is not necessary to find the optimal solution, although this will produce a deeper cut). Then the new row added to the tableau is

$$(24) \qquad -\alpha y - \beta v + v^{(k+1)} = -\nu$$

pivot operations are performed to locate the next extreme point (y_{k+1}, v_{k+1}) and so forth.

This algorithm is similar to that in [5] in that it is finitely convergent, cuts away an arbitrarily chosen extreme point at each step, and uses disjunctive dual programs to generate the cuts. The main difference is that the algorithm of Jeroslow generates a valid inequality for $E(j_k, Q_k)$ at each step, while our algorithm generates a valid inequality for $\text{conv}(Q_k \cap P(d_k-1))$. Indeed, the only use we make of the sets $E(j,P)$ is to test whether the current point is in the feasible set S. The linear program (23) compares unfavorably with the analogous program in [5, (7a-7b)] if the sets $E(j,P)$ are simple, i.e., the union of small numbers of faces. In the case of complex $E(j,P)$ our program (23), which completely avoids the E-sets, may be superior. In both cases the LPs are not as large as they look, and computational tricks and simplifications will be investigated later.

<u>An Example of Non-Convergence</u>:

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 R^3 be the polytope whose extreme points are

(25)	(3,3,0)	(3,3,5)	
	(0,8,0)	(0,8,θ)	[θ and φ to be specified later]
	(8,0,0)	(8,0,θ)	
	(3,10,0)	(3,10,φ)	
	(10,3,0)	(10,3,φ)	
	(8,8,0)		

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

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

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

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

Let θ,φ satisfy

(27) $$4 < \theta < 5 \quad \phi = \frac{3}{8}\theta$$

Then the extreme point (0,8,θ) can be cut away by the inequality

(28) $$(5-\phi)x + (5-\phi)y + 7z \leq 65 - 6\phi$$

(28) is a facet of E(2,P) which goes through (3,3,5); (3,10,φ); and (10,3,φ).

$Q_1 = P \quad (28)$ has the same extreme points as $Q_o = P$ except that $(0,8,\theta)$ and $(8,0,\theta)$ are replaced by $(0,8,\theta')$; $(8,0,\theta')$ such that

(29) $$\phi > \frac{3}{8}\theta' \qquad \theta' = \frac{2}{7}\phi + \frac{25}{7}$$

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

(30) $$\theta' x + \theta' y + 8z \leq 16\theta'$$

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

$Q_2 = Q_1 \quad (30)$ has the extreme points $(10,3,\phi)$ $(3,10,\phi)$ replaced by $(10,3,\phi')$ and $(3,10,\phi')$ where

(31) $$4 < \theta' < 5 \qquad \phi' = \frac{3}{8}\theta'$$

Since (31) is the same as (27) 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 in any limiting sense.

Concluding Remarks

There are two areas in need of investigation. Firstly, the behavior of computer implementations of these ideas needs to be studied (as previously stated, this paper concentrates on theoretical issues and an actual implementation would depend on further tricks). Secondly, 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 the ones specified are used (this is related to the creation of unwanted extreme points in our example). The author feels that present convergence proofs are more cumbersome than they should be, and that a theory unifying the various techniques is needed. The idea behind all convergence proofs for cutting-plane methods is that the polytope after a cut is "simpler" than the polytope before the cut. We lack a thorough understanding of what constitutes appropriate definitions of "simpler."

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. Balas, Egon. "Disjunctive Programming: Properties of the Convex Hull of the Feasible Points," MSRR No. 348, GSIA, Carnegie-Mellon University. (1974)

2. Balas, Egon. "Disjunctive Programming," MSRR No. 415. Presented at Discrete Optimization Conference in Vancouver, 1977.

3. Blair, Charles and Jeroslow, Robert. "A Converse for Disjunctive Constraints." Journal of Optimization Theory 25 (1978), pp. 195-206.

4. Gomory, Ralph. "An Algorithm for Integer Solutions to Linear Programming." IBM Technical Report 1958.

5. Jeroslow, Robert. "A Cutting-Plane Game and Its Applications," CORE discussion paper no. 7724. (1977)