

IMPROVED BOUNDS FOR INTEGER PROGRAMS:
A SUPERGROUP APPROACH

David E. Bell

November 1973

Research Memoranda are informal publications relating to ongoing or projected areas of research at IIASA. The views expressed are those of the author, and do not necessarily reflect those of IIASA.

Improved Bounds for Integer Programs:

A Supergroup Approach

David E. Bell*

Abstract

Recent work by Fisher and Shapiro has used Gomory's group theoretic methods together with Lagrange multipliers to obtain bounds for the optimal value of integer programs. Here it is shown how an extension of the associated abelian group to a supergroup can improve the bound without the artificiality of adding a cut. It is proved that there exists a finite group for which the dual ascent procedure of Fisher and Shapiro converges to the optimal integer solution. Constructive methods are given for finding this group. A worked example is included.

A primal-dual ascent algorithm of Fisher and Shapiro [3] uses Lagrange multipliers with Gomory's group problems [5] to give bounds on the optimal value of integer programs. If the ascent procedure does not discover the optimal solution, a cut is available to add to the original I.P. problem after which the whole process of forming the group problem and applying the ascent procedure may be repeated. This paper gives a method of improving the bound without the need of adding a cut by extending the group to a supergroup. It is proved that a finite supergroup exists for which the ascent procedure gives the optimal solution, and a constructive method for finding this group is given.

The first section outlines the basic ideas of the primal-dual ascent method of Fisher and Shapiro.

* This research was part of a doctoral thesis supervised by Professor J.F. Shapiro at the Operations Research Center, Massachusetts Institute of Technology, U.S.A.

1. The Primal-Dual Ascent Algorithm

The general linear integer programming problem is

$$\begin{aligned} \min \quad & \bar{c} w \\ \text{Aw} = & \bar{b} \\ \text{w} \geq & 0 \text{ integer,} \end{aligned} \tag{1}$$

which may be rewritten in terms of a dual feasible L.P. basis B as

$$\begin{aligned} z^* = \min \quad & cx \\ \text{s.t.} \quad & Nx \leq b \\ & Nx \equiv b \\ & x \geq 0 \text{ integer,} \end{aligned} \tag{2}$$

where $A = (B, \bar{N})$, $w = (x_B, x)$, $b = B^{-1}\bar{b}$, $N = B^{-1}\bar{N}$ and $c \geq 0$ is the modified cost vector. The symbol ' \equiv ' will represent equality with respect to addition modulo 1 unless indicated otherwise.

Define

$$L(u, x) = cx + u(Nx - b)$$

for non-negative vectors u , and call

$$L(u) = \min_{x \in X} L(u, x) \tag{3}$$

the Lagrangian for problem (2), where

$$X = \{x \geq 0 \mid Nx \equiv b, x \text{ integral}\} .$$

Problem (3) is a shortest route problem and is easily solved [6]. The following set definitions will be useful.

$$X(u) = \{x \in X \mid L(u, x) = L(u)\}$$
$$T(u) = \{t \in T \mid x^t \in X(u)\} ,$$

where T is an index set for X .

The following properties of the Lagrangian are well known [1, 3].

Lemma 1. For $u \geq 0$, $c + uN \geq 0$, the Lagrangian is continuous, concave, and a lower bound for the optimal I.P. value.

Since $L(u)$ is a lower bound for the I.P. the best of these bounds is evidently given by the dual problem

$$L = \max_{u \geq 0} L(u) . \quad (4)$$

To solve (4) consider the following reformulation

$$L = \max w$$
$$\text{s.t. } w \leq L(u, x) \quad x \in X \quad (5)$$
$$u \geq 0$$

or, equivalently,

$$L = \min \sum_{t \in T} \lambda_t (cx^t)$$
$$\sum_{t \in T} \lambda_t (Nx^t - b) \leq 0 \quad (6)$$
$$\sum_{t \in T} \lambda_t = 1$$
$$\lambda_t \geq 0 .$$

This problem is straightforward to solve using column generation but this approach has been found to be slow and so an ascent procedure for $L(u)$ was devised which gives monotonically increasing lower bounds. The idea is to generate, from any given vector \bar{u} a new vector u^* , for which $L(u^*) > L(\bar{u})$.

It can be shown [1, 3] that if $s^* > 0$ the optimal value of v, v^* say, in the linear program (7), gives a direction of ascent from \bar{u} , that is, for some $k > 0$,

$$L(\bar{u} + kv^*) > L(\bar{u})$$

and if $s^* = 0$ then \bar{u} is optimal.

$$\begin{aligned} s^* &= \max s \\ s &\leq 1 \\ s &\leq v(Nx - b) && x \in X(\bar{u}) \\ v_j &\geq 0 && i \in I(\bar{u}) \\ va_j &\geq 0 && j \in J(\bar{u}) \end{aligned} \tag{7}$$

where $I(u) = \{i | u_i = 0\}$, $J(u) = \{j | c_j + ua_j = 0\}$ and s has been arbitrarily bounded above by 1.

If $s^* > 0$ the new vector u^* is chosen by setting $u^* = \bar{u} + k^0 v^*$ where k^0 is the maximum value such that

$$L(\bar{u} + k^0 v^*) = L(\bar{u}) + k^0 s^* .$$

When the optimal value $L = L(u^*)$ has been found, if the optimal solution, λ^* , to (6) is integral, that is $\lambda_t^* = 0$ for all but one $x \in X$, then that x must be the optimal solution. If λ^* is not integral then the cut

$$L(u^*, x) \geq L(u^*) = L$$

may be added to (1) and the whole procedure repeated, starting from a dual feasible L.P. basis for the new problem.

The aim here is to provide a means of improving the bound L without the necessity of adding a cut.

2. The Method

The I.P. problem (2) may be written as

$$\begin{aligned} \min \quad & cx \\ \text{s.t.} \quad & x_B + Nx = b \\ & x_B, x \geq 0 \text{ integral,} \end{aligned} \tag{8}$$

where x_B are the basic variables.

Define variables y, \bar{y} by the relation

$$x_B = \Delta \bar{y} + y,$$

where $\Delta = \begin{pmatrix} \delta_1 & & 0 \\ & \ddots & \\ 0 & & \delta_m \end{pmatrix}$ is a diagonal matrix of

positive integers and $0 \leq y < \delta$. Since the condition $x_B \geq 0$ is implied by the two conditions $\bar{y}, y \geq 0$, (8) may be written equivalently as

$$\begin{aligned} \min \quad & cx \\ \text{s.t.} \quad & y + Nx \leq b \\ & y + Nx \equiv b \pmod{\delta} \\ & y, x \geq 0 \text{ integer.} \end{aligned} \tag{9}$$

Relaxing the conditions $\bar{y} \geq 0$ leaves the following problem in terms of y, x ,

$$\begin{aligned} \min \quad & cx \\ & y + Nx \equiv b \pmod{\delta} \\ & y, x \geq 0 \text{ integer.} \end{aligned} \tag{10}$$

Note that Gomory's asymptotic problem is just (10) with $\delta_i = 1, i = 1, \dots, m$. Once again (10) is a shortest route problem and is easily solved.

Let G be the abelian group in the original formulation and G^* that generated in (10).

Theorem 2. $G^* \cong G \oplus Z_1 \oplus \dots \oplus Z_m$;

that is, G^* is isomorphic to the direct sum of G with m cyclic groups, where Z_i has order δ_i .

Proof. Consider the set

$$S = \{s : s = g + \sum_{i=1}^m \mu_i e_i \quad 0 \leq \mu_i < \delta_i, g \in G \\ 0 \leq g < 1\} .$$

The order of S is $|G| \prod_{i=1}^m \delta_i$, for if

$$g^1 + \sum \mu_i^1 e_i = g^2 + \sum \mu_i^2 e_i ,$$

then

$$g^1 \equiv g^2 \pmod{1}$$

so that

$$\mu_i^1 \equiv \mu_i^2 \pmod{\delta_i} \text{ for all } i$$

and hence

$$g^1 = g^2, \mu_i^1 = \mu_i^2 .$$

It remains to show that S and G^* are identical.

G^* is generated by $\{e_1, \dots, e_m, a_1, \dots, a_n\}$ so let

$$g^* \equiv \sum_{j=1}^n \lambda_j a_j + \sum_{i=1}^m \mu_i e_i \pmod{\delta}$$

with μ, λ both integral.

Now

$$\sum_{j=1}^n \lambda_j a_j \equiv g \quad \text{for some } g \in G,$$

hence

$$\sum_{j=1}^n \lambda_j a_j = g + \sum \overline{\mu_i} e_i$$

so that

$$g^* = g + \sum_{i=1}^m (\mu_i + \overline{\mu_i}) e_i \quad (\text{mod } \delta).$$

Hence if

$$\mu^* \equiv \mu + \overline{\mu} \quad (\text{mod } \delta), \quad 0 \leq \mu^* < \delta,$$

then

$$g^* = g + \sum_{i=1}^m \mu_i^* e_i$$

and $g^* \in S$. Thus

$$G^* \subseteq S.$$

If $s \in S$ and

$$s = g + \sum_{i=1}^m \mu_i e_i,$$

where

$$g \equiv \sum_{j=1}^n \lambda_j a_j \quad (\text{mod } 1)$$

or

$$g = \sum_{j=1}^n \overline{\lambda_j} a_j + \sum_{i=1}^m \overline{\mu_i} e_i,$$

then

$$s = \sum_{j=1}^n \overline{\lambda_j} a_j + \sum_{i=1}^m (\mu_i + \overline{\mu_i}) e_i$$

and

$$s \pmod{\delta} \in G^*.$$

But $0 \leq s < \delta$, hence $s \in G^*$ and thus $S \subseteq G^*$.

||

Corollary . $|G^*| = |G| \cdot |\det \Delta|$.

Proof . It follows directly from the theorem. It may easily be seen that $|G^*|$ divides $|\det B| \cdot |\det \Delta|$ since G^* is generated in the ordinary manner by taking \bar{y} as the basis for the group problem in the L.P.

$$\begin{aligned} & \min \quad cx \\ & B\bar{\Delta}y + By + \bar{N}x = \bar{b} \\ & \bar{y}, y, x \geq 0 \end{aligned} \tag{11}$$

and therefore (see [10]) $|G^*|$ divides $|\det B\Delta|$. ||

The derivation of the Smith Normal Form of BA which is required for the simplification of the group equations [7] does not seem to be made easier by knowing the corresponding form for B but for a certain case it is possible.

Theorem 3 . If q_1, \dots, q_r are the elementary divisors of B , and $\det B, \delta_1, \dots, \delta_m$ are each mutually coprime then the elementary divisors of BA are $q_1, \dots, q_{r-1}, |\det \Delta|q_r$.

Proof . Consider the matrix B^1 which is B with the first column multiplied by δ_1 . It is sufficient to show that q_{r-1} is unaltered. But q_{r-1} represents the greatest common divisor of the $(m-1) \times (m-1)$ minors of B , (see [8]). Now there are m such minors which are left constant in B^1 , that is they contain no elements of column 1. But $|\det B|$ is a linear combination of these m minors thus if they have a common factor it must divide $|\det B|$. But δ_1 and $|\det B|$ are coprime thus q_{r-1} is unchanged. By repetition with $\delta_2, \dots, \delta_m$ the theorem is proved. ||

The purpose of this new formulation is to increase further the lower bound L obtained by the primal-dual ascent method.

Let $X^* = \{(y,x) \geq 0 \mid y + Nx \equiv b \pmod{\delta}, y, x \text{ integral}\}$
with index set T^* .

Lemma 5. The index sets T, T^* are equivalent.

Proof. If $x \in X$ then $Nx - b$ is integral so that $y \equiv b - Nx \pmod{\delta}$
is integral. Hence

$$(y,x) \in X^* .$$

If $(y,x) \in X^*$ then

$$y \equiv b - Nx \pmod{\delta} ,$$

and since y is integral, so is

$$b - Nx .$$

Hence $x \in X$.

Thus T^* may be considered to be precisely T and further
reference to it as a distinct set will not be necessary.
Recall that λ^* is the optimal solution to

$$\begin{aligned} L = \min \quad & \sum_{t \in T} \lambda_t (cx^t) \\ \text{s.t.} \quad & \sum_{t \in T} \lambda_t (Nx^t - b) \leq 0 \\ & \sum_{t \in T} \lambda_t = 1 \\ & \lambda_t \geq 0 \end{aligned} \tag{12}$$

with optimal dual vector u^* .

The Lagrangian associated with problem (10) is

$$L^*(u) = \min_{(y,x) \in X^*} \{cx + u(y + Nx - b)\}$$

yielding a corresponding primal problem

$$\begin{aligned}
 L^* &= \min \sum_{t \in T} \lambda_t (cx^t) \\
 \text{s.t.} \quad & \sum_{t \in T} \lambda_t (Nx^t + y^t - b) \leq 0 \\
 & \sum_{t \in T} \lambda_t = 1 \\
 & \lambda_t \geq 0 .
 \end{aligned} \tag{13}$$

The following theorem establishes some properties of (13).

- Theorem 6.
- i) $L \leq L^* \leq z^*$
 - ii) λ^* is infeasible if $\lambda_t^*(u^*y^t) \neq 0$ for some t
 - iii) $L < L^*$ if $u^*y^t > 0$ for all $t \in T(u^*)$.

Proof. i) Since $y^t \geq 0$ for all $t \in T$ any feasible solution of (13) is feasible in (12). Hence $L \leq L^*$. If x^k is the optimal integral solution, then $\lambda_k = 1$ is feasible in (13) since $Nx^k + y^k \leq b$, so that $L^* \leq z^*$.

$$\begin{aligned}
 \text{ii)} \quad & \sum_{t \in T} \lambda_t^* (N_i x^t + y_i^t - b_i) \\
 &= \sum_{t \in T} \lambda_t^* (N_i x^t - b_i) + \sum_{t \in T} \lambda_t^* y_i^t \\
 &= \sum_{t \in T} \lambda_t^* y_i^t \quad \text{if } u_i^* > 0 .
 \end{aligned}$$

If $\lambda_t^*(u^*y^t) \neq 0$ then

$$\sum_{t \in T} \lambda_t^* y_i^t > 0 \quad \text{for some } u_i^* > 0$$

so that λ^* is infeasible in (13).

iii) It is sufficient to show that

$$L^*(u^*) > L(u^*) .$$

If $x \in X(u^*)$ then

$$(c + u^*N)x + u^*y - u^*b > L(u^*)$$

since $u^*y > 0$ by assumption.

If $x \notin X(u^*)$ then it is easily shown by induction that u^* is rational and hence so is $(c + u^*N)x$ and thus

$$0 < \inf_{x \in X - X(u^*)} (c + u^*N)x - u^*b .$$

Hence

$$L^*(u^*) > L(u^*)$$

as required. ||

The algorithmic question is thus, in knowing the solution, λ^* , to problem (6), how should the value of δ be chosen?

For some $i \in I(u^*)$ let d_i be the greatest common divisor of $N_i x^t - b$ for those t for which $\lambda_t^* > 0$.

Corollary 6.1. λ^* is infeasible in (13) if δ_i does not divide d_i .

Proof. If $\delta_i \nmid d_i$ then $N_i x^t \not\equiv b_i \pmod{\delta_i}$ for some t for which $\lambda_t^* > 0$ so that the corresponding $y_i^t > 0$. Hence $\lambda_t^* u^* y_i^t > 0$ and by the theorem, λ^* is infeasible. ||

Note that the corollary includes the case $\delta_i > d_i$.

Corollary 6.2. If δ_i does not divide $N_i x^t - b_i$ for all $t \in T(u^*)$ for some $i \in I(u^*)$ then

$$L < L^* .$$

Proof. In this case, $u^* y_i^t > 0$ for all $t \in T(u^*)$. ||

For illustrative purposes consider the case when the bound L cannot be improved, that is, when $L = z^*$.

i) λ^* is integral. In this case $\lambda_t^* > 0$ only for the optimal feasible solution x^* . If $u_1^* > 0$ then $N_1 x^* = b_1$ so that all values of δ_i divide $N_1 x^* - b_1$. Hence the conditions of the corollaries cannot be met.

ii) λ^* is not integral. Then

$$z^* = cx^* \geq cx^* + u^*(Nx^* - b) = L(u^*, x^*) \geq L(u^*) = L.$$

Therefore $x^* \in X(u^*)$ and so the conditions of the second corollary cannot be met.

These sufficient conditions on Δ such that the bound may be improved are not very stringent and Δ might often be chosen so that $|\det \Delta|$ is not large, but this is a matter of tradeoff between efficiency and gain. Note that if $i \in I(u^*)$ it might as well be that $\delta_i = 1$ since in this case δ_i plays no part in the conditions.

The quantities $N_1 x^t - b_1$ are available in the optimal basis matrix Q of problem (13). This does not furnish all of $X(u^*)$ but is probably sufficient to choose Δ in practice.

Corollary 6.3. If p is the smallest prime not dividing $\det Q$ then if, for some $i \notin I(u^*)$, $\delta_i = p$, $\delta_j = 1$, $j \neq i$, then λ^* is infeasible.

Proof. In this case δ_i does not divide d_i since d_i divides $\det Q$. ||

3. An Example

Consider the Integer Program

$$\begin{aligned} \min \quad & 4x + y \\ \text{s.t.} \quad & 3x + y + 9w = 11 \\ & -12x + 7y + 3z = 11 \\ & w, x, y, z \geq 0 \text{ integer.} \end{aligned}$$

The optimal basic variables are (w, z) and the primal-dual ascent algorithm gives $L = 3$, $I(u^*) = \{1\}$. The active elements of X are $x^1 = (0, 2)$, $x^2 = (3, 2)$, $x^3 = (2, 5)$, the primal-dual solution being $\lambda^* = \left(\frac{11}{12}, \frac{1}{12}, 0\right)$.

$$Nx^1 - b = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad Nx^2 - b = \begin{pmatrix} 0 \\ -11 \end{pmatrix}$$

so that $d_1 = d_2 = 1$ and there are no restrictions on the choice of Δ . Since $u_1^* = 0$, δ_1 will be assigned the value 1.

The following table gives values of L^* for increasing values of δ_2 . The optimal solution to the problem is x^3 , hence $z^* = 13$.

δ_2	λ^*	L^*
1	$\frac{11}{12}, \frac{1}{12}, 0$	3
2	$\frac{5}{6}, \frac{1}{6}, 0$	4
3	$\frac{3}{4}, \frac{1}{4}, 0$	5
4	$\frac{2}{3}, \frac{1}{3}, 0$	6
5	$\frac{2}{3}, \frac{1}{3}, 0$	6
6	$\frac{1}{2}, \frac{1}{2}, 0$	8
7	$\frac{2}{3}, \frac{1}{3}, 0$	6
8	$\frac{1}{2}, \frac{1}{2}, 0$	8
9	$\frac{1}{3}, \frac{2}{3}, 0$	10
10	$\frac{1}{2}, \frac{1}{2}, 0$	8
11	$\frac{1}{2}, \frac{1}{2}, 0$	8
12	0, 0, 1	13

at which point the gap is closed.

4. Convergence

The example had the pleasing property that for some value of δ , namely (1, 12), the primal-dual ascent method gave the optimal I.P. solution.

Theorem 7. If the L.P. feasible region of the problem is bounded, then there exists a finite vector δ for which the modified ascent method gives the optimal integral solution, or shows that there is none.

Proof. Since the L.P. region is bounded, there exists some finite $\alpha, \beta > 0$ such that

$$\begin{aligned} \beta &> \max \quad \alpha x \\ &\text{s.t.} \quad Nx \leq b \\ &\quad \quad \quad x \geq 0 . \end{aligned}$$

Let

$$Y = \{x \geq 0 \mid \alpha x \leq \beta\}$$

and define

$$D_i = \max \left\{ \max_{x \in Y} (N_i x - b_i) , -\min_{x \in Y} (N_i x - b_i) \right\}$$

and since D_i is finite, let δ_i be the smallest finite integer strictly greater than D_i , for all $i = 1, \dots, m$. Note then that for all $x^t \in X \cap Y$ if

$$\begin{aligned} \text{then} \quad Nx^t + y^t - b &\equiv 0 \pmod{\delta} , \\ N_i x^t + y_i^t - b_i &= 0 \quad \text{if } N_i x^t \leq b_i , \\ N_i x^t + y_i^t - b_i &= \delta_i \quad \text{if } N_i x^t > b_i . \end{aligned}$$

So if

$$x^t \in X \cap Y$$

then

$$Nx^t + y^t - b \geq 0 .$$

Consider the solution to

$$\begin{aligned}
 \min \quad & \sum_{t \in T} \lambda_t (cx^t) \\
 & \sum_{t \in T} \lambda_t (Nx^t + y^t - b) \leq 0 \\
 & \sum_{t \in T} \lambda_t = 1 \\
 & \lambda_t \geq 0 .
 \end{aligned} \tag{14}$$

Let

$$\begin{aligned}
 T^0 &= \{t \in T \mid Nx^t + y^t - b \geq 0\} \\
 \bar{T}^0 &= \{t \in T \mid Nx^t + y^t - b \not\geq 0\}
 \end{aligned}$$

and λ^* be the solution to (14).

Thus

$$\sum_{t \in T^0} \lambda_t^* (Nx^t + y^t - b) + \sum_{t \in \bar{T}^0} \lambda_t^* (Nx^t + y^t - b) \leq 0 ,$$

hence

$$\sum_{t \in \bar{T}^0} \lambda_t^* (Nx^t - b) \leq 0 .$$

If

$$\sum_{t \in \bar{T}^0} \lambda_t^* > 0 ,$$

consider the point

$$x^* = \frac{1}{\sum_{t \in \bar{T}^0} \lambda_t^*} \cdot \sum_{t \in \bar{T}^0} \lambda_t^* x^t .$$

Since $x^* \geq 0$, $Nx^* \leq b$, then $\alpha x^* < \beta$. But x^* is a convex combination of vectors satisfying $\alpha x > \beta$ (since all $x^t \in X \cap Y$ have $t \in T^0$) which is a contradiction. Hence

$$\lambda_t^* = 0 \text{ for all } t \in \bar{T}^0,$$

so that $\lambda_t^* > 0$ implies $t \in T^0$ and (14) forces

$$Nx^t + y^t - b = 0 ,$$

so that x^t is feasible and thus optimal.

Problem (14) can have no feasible solution if and only if there is no feasible I.P. solution. ||

Although the theorem only deals with bounded linear programs, clearly any unbounded L.P. with a known I.P. upper bound can be converted.

5. Summary

A number of conditions were given concerning a suitable choice of δ , depending on the optimal solution λ^* and the corresponding active elements of X . It is certainly not necessary to await optimality before choosing a value of δ and, of course, δ may be increased repeatedly as desired. Thus at any time there is available a choice of adding the cut

$$L(u, x) \geq L(u)$$

and restarting or choosing a higher value of δ . A third method, not discussed here, is to choose an alternative basis to problem (1), other than B , on which to apply the ascent algorithm (see Bell and Fisher [2]).

Preliminary computation with the ascent algorithm appears promising. This extension provides a simple constructive method for resolving any duality gap.

References

- [1] Bell, D.E. "The Resolution of Duality Gaps in Discrete Optimization," Technical Report No. 81, Operations Research Center, M.I.T., Cambridge, August 1973.
- [2] Bell, D.E. and Fisher, M.L. "Improved Integer Programming Bounds Using Intersections of Corner Polyhedra," University of Chicago Graduate School of Business Working Paper, Chicago, October 1973.
- [3] Fisher, M.L. and Shapiro, J.F. "Constructive Duality in Integer Programming," to appear in SIAM Journal on Applied Mathematics.
- [4] Geoffrion, A.M. "Lagrangian Relaxation and Its Uses in Integer Programming," Western Management Science Institute, University of California at Los Angeles Working Paper No. 195, Los Angeles, December 1972.
- [5] Gomory, R.E. "Some Polyhedra Related to Combinatorial Problems," Lin. Alg., 2 (1969), 451-558.
- [6] Gorry, G.A., Northup, W.D., and Shapiro, J.F. "Computational Experience with a Group Theoretic Integer Programming Algorithm," Math. Prog., 4 (1973), 171-192.
- [7] Gorry, G.A. and Shapiro, J.F. "An Adaptive Group Theoretic Algorithm for Integer Programming Problems," Man. Sci., 17 (1971), 285-306.
- [8] Mostow, G.D., Sampson, J.H., and Meyer, J.P. Fundamental Structures of Algebra. New York, McGraw-Hill, 1963.
- [9] Shapiro, J.F. "Generalized Lagrange Multipliers in Integer Programming," Operations Research, 19 (1971), 68-76.
- [10] Wolsey, L.A. "Group Representation Theory in Integer Programming," Technical Report No. 41, Operations Research Center, M.I.T., Cambridge, 1969.