AD-A036 860 CARNEGIE-MELLON UNIV PITTSBURGH PA MANAGEMENT SCIENC--ETC F/G 12/2
THE VALUE FUNCTION OF A MIXED INTEGER PROGRAM: II.(U)
DEC 76 C E BLAIR, R G JEROSLOW N00014-75-C-0621

UNCLASSIFIED RR-397 NL

| OF |

AD
A036860

END
DATE
FILMED
4-77

END
DATE
FILMED
4-77

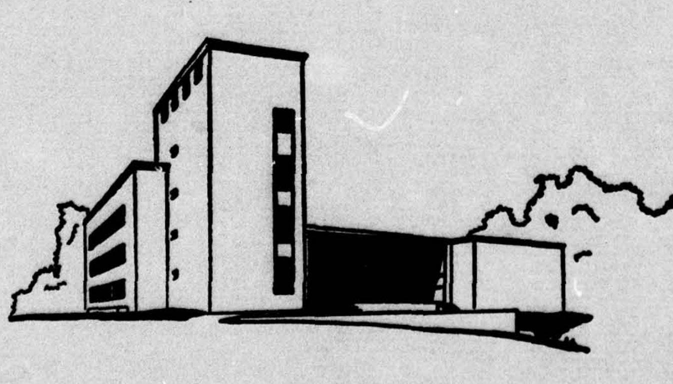
1

Carnegie-Mellon University

PITTSBURGH, PENNSYLVANIA 15213

GRADUATE SCHOOL OF INDUSTRIAL ADMINISTRATION

WILLIAM LARIMER MELLON, FOUNDER



Management Sciences Research Report No. 397

THE VALUE FUNCTION OF A
MIXED INTEGER PROGRAM: II

by

C.E.Blair and R.G. Jeroslow

December 1976

This report was prepared as part of the activities of the Management Sciences Research Group, Carnegie-Mellon University, under Grant NSF - #GP-37510 X1 of the National Science Foundation and Contract N00014-75-C-0621 NR047-048 with the U.S. Office of Naval Research. Reproduction in whole or in part is permitted for any purpose of the U.S. Government.

Management Science Research Group
Graduate School of Industrial Administration
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

See 1473

Abstract

We prove that the gap in optimal value, between a mixed-integer program in rationals and its corresponding linear programming relaxation, is bounded as the right-hand-side is varied. In addition, a variant of value iteration is shown to construct subadditive functions which resolve a pure-integer program when no dual degeneracy occurs. These subadditive functions provide solutions to subadditive dual programs for integer programs which are given here, and for which the values of primal and dual problems are equal.

Key words:

1) Integer programming
2) Cutting-planes
3) Convex analysis
4) Duality
5) Subadditivity

THE VALUE FUNCTION OF A
MIXED INTEGER PROGRAM: II

by

C.E. Blair and R.G. Jeroslow

In this paper, we continue our study, begun in [2], of the value function of a mixed-integer program. This function G provides the value G(b) of the program

$$\text{(MIP)} \qquad \begin{array}{rl} \inf & cx + dy \\ \text{subject to} & Ax + By = b \\ & x,\ y \geq 0 \\ & x \text{ integer} \end{array}$$

in variables $x = (x_1,\ldots,x_r)$, $y = (y_1,\ldots,y_s)$, with right-hand-side (r.h.s.) $b = (b_1,\ldots,b_m)$ and matrices A, B and vectors c,d conformally dimensioned. G(b) is defined only for b feasible, i.e., b for which the constraints of (MIP) are consistent.

In [2], we emphasized bounds on G(b) [2, Theorem 2.1], a structure theorem for G(b) [2, Theorem 3.3], and the extendability of G(b) [2, Theorem 4.6]. Here our emphasis is on the difference between G(b) and the value function L(b) of the linear programming relaxation (LP) of (MIP) (see (LP) below). We show that the gap in value is finite (Corollary 1.3) and that in certain regions $S_{\bar{\gamma}}$ (defined below) the gap function is subadditive (Theorem 2.6). Our investigation of the gap function leads quite naturally to a subadditive dual program for (MIP) (Theorem 2.4), which places the subadditive functions into a more tractable class of functions than does an earlier subadditive dual (see [6]).

We also show that the subadditive functions needed in our dual programs can all be obtained by value-iteration, which replaces a subadditive function F_n by an improved function F_{n+1} that provides a better value in the dual program, with finite convergence guaranteed under many commonly-occurring hypotheses (Theorem 2.8).

Throughout our work, we have a standing assumption:

(SA) A, B, b are matrices of rationals;

which was utilized also in [2]. For the pure-integer case (s = 0), this assumption (SA) takes the equivalent form

(SA)' A,b are matrices of integers

In addition, we assume throughout that G(0) = 0 and $r \geq 1$ if s = 0.

The assumption G(0) = 0 is known to imply that (LP) below has a finite value if it is feasible (see e.g., [6]). The assumptions made here will be utilized below without explicit citation.

In what follows, we also use the notation $A = [a^{(j)}]$ (cols) and $B = [b^{(k)}]$ (cols).

Section 1: Value Gaps and Quadratic Duals

We begin by stating a result (Theorem 1.2) which is useful in what follows. The result depends on [2, Theorem 2.1 (2)] and requires the following lemma, which is implicit in [7, Lemma 1]. We provide a self-contained proof here, since it is very short; for a generalization of the lemma below to multicriteria objective functions, see [7].

Lemma 1.1: Suppose that z^o is an optimum for the linear program

$$\text{(1)} \qquad \begin{array}{rl} \min & dz \\ \text{subject to} & Dz = h \\ & z \geq 0 \qquad (z = (z_1,\dots,z_p)) \end{array}$$

and $z^* \geq 0$ satisfies

(2) $\quad z_j^* > 0$ implies $z_j^o > 0$ for $j = 1,\dots,p$

Then if $h' = Dz^*$, z^* is optimal for the linear program

$$\text{(1)'} \qquad \begin{array}{rl} \min & dz \\ \text{subject to} & Dz = h' \\ & z \geq 0 \end{array}$$

In addition, an optimum in the dual to (1) is also optimum in the dual to (1)'.

Proof: The program (1) has the dual

$$\text{(3)} \qquad \begin{array}{rl} \max & \lambda h \\ \text{subject to} & \lambda D \leq d \end{array}$$

Let λ^o be an optimum to (3), so that the following complementary slackness condition holds:

(4) $$z^o(d - \lambda^o D) = 0$$

From (2), (4) we obtain this version of complementary slackness:

(4)' $$z^*(d - \lambda^o D) = 0 .$$

Now (4)' is widely known to imply that z^* is optimal in (1)' . In detail, from (4)' $dz^* = \lambda^o Dz^* = \lambda^o h'$, showing that z^* is optimal in (1)' and λ^o is optimal in the dual to (1)'.

Q.E.D.

<u>Theorem 1.2</u>:

There is a constant $\Delta \geq 0$, independent of b in (MIP), which possesses the following property:

If (x^o, y^o) is an optimum to the linear programming relaxation

$$\text{(LP)} \qquad \begin{aligned} \min \quad & cx + dy \\ \text{subject to} \quad & Ax + By = b \\ & x, y \geq 0 \end{aligned}$$

and (MIP) is consistent, there is an optimum (x^I, y^I) to (MIP) with

(5) $$| (x^I, y^I) - (x^o, y^o) | \leq \Delta$$

<u>Proof</u>: Let x^* be the vector of "integral parts" of x^o, i.e, $x_j^* = \lfloor x_j^o \rfloor$ for $j = 1,\ldots,r$. Put $b' = Ax^* + By^o$. By Lemma 1, (x^*, y^o) is optimal in the program

$$\text{(6)} \qquad \begin{aligned} \min \quad & cx + dy \\ \text{subject to} \quad & Ax + By = b' \\ & x, y \geq 0 \\ & x \text{ integer} \end{aligned}$$

since it is optimal when the requirement "x integer" is dropped in (6), and yet x^* does satisfy the integrality requirement. From [2 , Theorem 2.1 (1)] there is a solution (x^I,y^I) to (MIP) with

$$(7) \qquad |\, (x^I, y^I) - (x^*,y^o) \,| \leq C \,|\, b - b' \,| + D$$

with $C,D \geq 0$ independent of b,b' . Using (7), we have

$$(8) \qquad |\, (x^I, y^I) - (x^o,y^o) \,| \leq |\, (x^I,y^I) - (x^*,y^o) \,| + |\, (x^*, y^o) - (x^o, y^o) \,|$$

$$< C \,|\, \Sigma \,(x_j^o - \lfloor x_j^o \rfloor)\, a^{(j)} \,| + D + r \quad .$$

$$< C \left(\sum_{j=1}^{r} |\, a^{(j)} \,| \right) + D + r$$

By taking Δ to be the r.h.s. of (8), which is clearly independent of b, the Theorem is immediate.

Q.E.D.

In what follows, L(b) is the value function of the linear relaxation of (MIP), i.e., L(b) is the value of the linear program (LP) as a function of the r.h.s. b. Clearly, $G(b) \geq L(b)$ when b is a feasible r.h.s. for (MIP).

<u>Corollary 1.3</u>:

$$(9) \qquad \sup \{\, |\, G(b) - L(b) \,| \;\; |\, b \text{ feasible in (MIP)} \,\} < +\infty$$

<u>**Proof**</u>: From (5) and Theorem 1.2,

$$|\, G(b) - L(b) = |\, c\,(x^I - x^o) + c\,(y^I - y^o) \,|$$

$$\leq |\, c\,|\,|x^I - x^o| + |d|\,|y^I - y^o| \leq (|c| + |d|)\,\Delta < +\infty$$

Q.E.D.

It is worth noting that Corollary 1.3 implies

$$\left| G(b) - G(b') \right| \leq \left| G(b) - L(b) \right| + \left| L(b) - L(b') \right| + \left| L(b') - G(b') \right|$$

$$\leq K \left| b - b' \right| + 2\,\delta,$$

where K is any constant for which

$$\left| L(b) - L(b') \right| \leq K \left| b - b' \right|,$$

δ is the supremum defined in (9), and b,b' are feasible. This improves [2, Theorem 2.1 (2)], in that the modulus associated with $|b-b'|$ can be taken to be that of (LP), if one is willing to use $2\,\delta$ as a constant. Note, however, that in a pure-integer program (s = 0) it may still be necessary to retain a positive constant, if the modulus used is K of the linear program (LP). A constant of zero can be employed for s = 0, if the modulus is $K + 2\,\delta$, since $|b-b'| \neq 0$ implies $|b-b'| \geq 1$ for integer vectors b,b'.

For a function f: $S \to R$, S any set, we define the epigraph epi(f) by:

(10) $\quad \mathrm{epi}(f) = \{ (y,x) \mid x \in S \text{ and } y \geq f(x) \}$

For a set $T \subseteq R^n$, conv (T) respectively clconv (T) shall denote the convex span resp. closed convex span of T. Compare with [10], [13].

<u>Proposition 1.4</u>:

(11) $$\mathrm{epi}(L) = \mathrm{conv}(\mathrm{epi}(G))$$

<u>Proof</u>: Since $L(b) \leq G(b)$ for b in the domain of G, and epi(L) is convex due to the convexity of the domain of L, the inclusion ($\supseteq$) in (11) is clear.

To establish the inclusion ($\subseteq$) in (11), and thereby complete the proof, let $(z,b) \in \text{epi}(L)$, so that $z \geq G(b)$. Let (x^o, y^o) be an optimum to (LP), so that $G(b) = cx^o + dy^o$. When b is rational, we can assume (x^o, y^o) is rational by (SA), so that $(x^o, y^o) = (x^*/D, y^o)$ with $x^* \geq 0$ a vector of non-negative integers, and $D \geq 1$ an integer.

Since $L(Db) = DL(b)$, as one easily proves using Lemma 1.1, (x^*, y^oD) is optimal to (LP) with r.h.s. Db, and since x^* is integral, we see that (x^*, y^oD) is also optimal to (MIP) with r.h.s. Db. Then $G(Db) = x^* + y^oD = L(Db) = DL(b) \leq D\,z$, giving $(Dz, Db) \in \text{epi}(G)$. Since $(0,0) \in \text{epi}(G)$,

$(Dz, Db)/D + (0,0)\,(D-1)/D = (z,b) \in \text{conv}(\text{epi}(G))$.

For b irrational and $(z,b) \in \text{epi}(L)$, we also obtain $(z,b) \in \text{epi}(G)$ by the continuity of L and the fact that b can be approximated by rational r.h.s.

Q.E.D.

We next present a result which, in Rockafellar's duality framework [11], [12], can be construed as a dual program for the pure integer program (IP), which is (MIP) for $s = 0$:

$$\text{(IP)} \qquad \begin{array}{ll} \min & cx \\ \text{subject to} & Ax = b \\ & x \geq 0 \text{ and integer} \end{array}$$

Let $\alpha^{(i)}$ denote the i-th row of A and let b_i denote the i-th component of b. The quadratic Lagrangean for (IP) is defined to be

(12) $$L(x,\lambda,\rho) = (c - \lambda A)x + \rho \sum_{i=1}^{m} (\alpha^i x - b_i)^2 + \lambda b$$

and a dual problem for (IP) is

$(QD)_\lambda$ $$\max_{\rho \geq 0} \inf_{\substack{x \geq 0 \\ x \text{ integer}}} L(x,\lambda,\rho)$$

where $\lambda = (\lambda_1,\ldots,\lambda_m) \in R^m$ is a fixed vector.

Denote the value of (IP) by v(IP) and let $v(QD_\lambda)$ denote the quantity defined in $(QD)_\lambda$. It is immediate that $v(QD_\lambda) \leq v(IP)$, since for any ρ and any λ

(13) $$\inf_{\substack{x \geq 0 \\ x \text{ integer}}} L(x,\lambda,\rho) \leq \inf \{L(x,\lambda,\rho) \mid Ax = b,\ x \geq 0 \text{ and integer}\}$$

$$= \inf \{cx \mid Ax = b,\ x \geq 0 \text{ and integer}\} = v(IP)$$

<u>Theorem 1.5</u>: Suppose that (IP) is consistent.

(1) For any $\lambda \in R^m$ there is $\rho(\lambda) \geq 0$ such that for $\rho \geq \rho(\lambda)$ we have

(14) $$L(x^*,\lambda,\rho) = \min_{\substack{x \geq 0 \\ x \text{ integer}}} L(x,\lambda,\rho)$$

if and only if x^* is optimal in (IP). Furthermore, $\rho(\lambda)$ is independent of the r.h.s. b in (IP). In particular, $v(IP) = v(QD_\lambda)$.

(2) Let $\bar{\lambda}$ be an optimum to the linear dual

(15) $$\max \lambda b \quad \text{subject to } \lambda A \leq c$$

of the linear programming relaxation (LP) of (IP). Then if $\rho > G(b) - L(b)$, (14) for $\lambda = \bar{\lambda}$ holds precisely if x^* is optimal in (IP).

Proof:

(1) From [2, Theorem 2.1(2)], there is a constant $\epsilon \geq 0$, independent of b, for which

(16) $$| G(b) - G(b') | \leq \epsilon | b - b' |$$

when b' is a feasible r.h.s. in (IP).

Note that, if $b = (b_1,\ldots,b_m)$ and $b' = (b_1',\ldots,b_m')$ are integer vectors then

(17) $$| b - b' | \leq \sum_{i=1}^{m} (b_i - b_i')^2$$

where $|b - b'|$ denotes the norm used in [2], specifically (see [2 , eqn (1.1)])

(18) $$| b - b' | = \sum_{i=1}^{m} | b_i - b_i' |$$

Put $\rho(\lambda) = 1 + \epsilon + | \lambda |$. Then from (16) and (17), we have, for any vector $x \geq 0$ of integers and $\rho \geq \rho(\lambda)$,

(19) $$\begin{aligned} L(x,\lambda,\rho) &= cx + \lambda(b - Ax) + \rho \sum_{i=1}^{m} (\alpha^i x - b_i)^2 \\ &\geq G(b') + \lambda(b - b') + \rho | b - b' | \\ &\geq G(b) - \epsilon |b - b'| - |\lambda| |b - b'| + \rho |b - b'| \\ &> G(b) \end{aligned}$$

if $b' \neq b$, where $b' = Ax$; while if $b' = b$, we have

(20) $$L(x, \lambda, \rho) = cx \geq G(b)$$

Now a non-negative integer vector x^* is optimal in (IP) if and only if $b' = b$ and $cx^* = G(b)$. Hence by (20) if x^* is optimal, $L(x,\lambda,\rho) = G(b)$; while if x^* is not optimal, we have $L(x,\lambda,\rho) > G(b)$ by (19), (20). Therefore, (14) holds if and only if x^* is optimal in (IP), as claimed.

Regarding the "particular," from (19) and (20) we have $\inf\{L(x,\lambda,\rho) \mid x \geq 0 \text{ and integer}\} = G(b) = v(IP)$ by (19), (20). Since $v(QD_\lambda) \leq v(IP)$, this shows that $v(QD_\lambda) = v(IP)$.

(2) Suppose that $\rho > G(b) - L(b)$. Let $x \geq 0$ be an integer vector. If $b' = Ax$ is different from b, we obtain

(21) $$\begin{aligned} L(x, \bar{\lambda},\rho) &= (c - \bar{\lambda} A)x + \rho \Sigma (b_i - b_i')^2 + \bar{\lambda} b \\ &\geq 0 + \rho + L(b) \\ &> G(b) \end{aligned}$$

(In (21), we have used the fact that $c - \bar{\lambda}A \geq 0$ and $\bar{\lambda}b = L(b)$). If $b' = b$, we obtain (20).

Using (20),(21) as we used (19), (20) in proving the first part, we establish the present claim.

Q.E.D.

Rockafellar obtained a result of the type Theorem 1.5(1) for convex programs (see [12,Theorem 3.5]). Our result is of course easier to prove if the vector x is bivalent or bounded in (IP).

Theorem 1.5(2) is also more elementary than Theorem 1.5(1), since it does not require the proximity results of [2], while still indicating a choice $\bar{\lambda}$ of λ for which $\rho(\lambda)$ is likely to be smaller than for general λ.

As a consequence of Theorem 1.5, the value of the "quadratic dual"

$$\text{(QD)} \qquad \max_{\lambda \in R^m} \ \max_{\rho \geq 0} \ \inf_{\substack{x \geq 0 \\ x \text{ integer}}} L(x,\lambda,\rho)$$

is v(IP). Here the analogy to ordinary linear programming duality is easier to see, since fixing $\rho = 0$ one easily establishes that the program

$$\text{(22)} \qquad \max_{\lambda \in R^m} \ \inf_{\substack{x \geq 0 \\ x \text{ integer}}} L(x,\lambda,0)$$

$$= \max_{\lambda \in R^m} \ \inf_{x \geq 0} L(x,\lambda,0)$$

is the usual linear programming dual to (LP) for $s = 0$. However, since λ can be fixed in (QD) while still having $v(IP) = v(QD_\lambda)$, the dual (QD) actually has many features of a penalty method.

Section 2: Subadditive Dual Programs

We recall that a function F is subadditive on a domain S if

$$F(v + w) \leq F(v) + F(w) \tag{23}$$

is an identity for $v, w \in S$ (see [4], [5], [8]). Typically, S is required to have $0 \in S$ and to be closed under addition. In what follows, we take S to be the set of all feasible r.h.s. for (IP), which is readily verified to possess these properties. A particular class of subadditive functions are those which are non-negative and bounded:

$$\Sigma = \left\{ F \text{ is subadditive on } S \;\middle|\; \begin{array}{l} F(v) \geq 0 \text{ for all } v \in S \\ \text{and } \sup \{F(v) \mid v \in S\} < +\infty \end{array} \right\} \tag{24}$$

We shall also need the concept of the upper directional derivative (at zero) of a subadditive function F (see [5], [8]):

$$\overline{F}(v) = \lim\sup \{ F(\lambda v)/\lambda \mid \lambda \searrow 0^+ \} . \tag{25}$$

We recall two results from previous work; the first can be proven directly.

<u>Lemma 2.1</u>: (See [5]). If F_1 and F_2 are subadditive on S, so is $F_1 + F_2$. In particular, for any subadditive function F and any $\lambda \in R^m$, the function G defined by

$$G(v) = \lambda v + F(v) \tag{26}$$

is subadditive on S.

Lemma 2.2: (See [5]).

If G is subadditive on S, then

(27) $$\sum_{j=1}^{r} G(a^{(j)})x_j + \sum_{k=1}^{s} \overline{G}(b^{(k)})y_k \geq G(b)$$

for any feasible solution (x,y) to (MIP).

We shall also need one more lemma.

Lemma 2.3: If F is a non-negative, subadditive function on S, and $\delta \geq 0$ is any scalar, then the function H defined by

(28) $$H(v) = \min \{ F(v), \delta \}$$

is in the class Σ of (24).

Proof: Let $v,w \in S$. If $F(v) \geq \delta$ or if $F(w) \geq \delta$, (28) gives

(29) $$H(v + w) \leq \delta \leq H(v) + H(w)$$

as F is non-negative. If F(v) and F(w) are $< \delta$, then

(30) $$H(v + w) \leq F(v + w) \leq F(v) + F(w) = H(v) + H(w).$$

Since the two cases (29), (30) are exhaustive of the possibilities, H is subadditive. Clearly H is non-negative, and H is bounded by δ; hence $H \in \Sigma$.

Q.E.D.

Theorem 2.4: Suppose that (MIP) is consistent.

Let λ be any feasible solution for the dual

(DLP)
$$\begin{array}{rl} \max & \lambda b \\ \text{subject to} & \lambda A \leq c \\ & \lambda B \leq d \end{array}$$

to (LP). Then the subadditive program

$$(D_\lambda) \qquad \begin{array}{ll} \max & \lambda b + F(b) \\ & \lambda a^{(j)} + F(a^{(j)}) \leq c_j \ , \ j = 1,\ldots,r \\ & \lambda b^{(k)} + \overline{F}(b^{(k)}) \leq d_k \ , \ k = 1,\ldots,s \\ & F\epsilon\Sigma \end{array}$$

has the value v(MIP) of MIP, and moreover for any feasible solution (x,y) to (MIP) and any feasible solution F to (D_λ), we have (27) for G as defined in (26).

Proof: The bound (27) is simply a consequence of Lemmas 2.1 and 2.2. This bound also shows that $F(b) + \lambda b \leq v(MIP)$ for F feasible in (D_λ), as $G(a^{(j)}) \leq c_j$ and $\overline{G}(b^{(k)}) \leq d_k$ for all j,k.

It suffices to prove that there is a solution H to (D_λ) with $\lambda b + H(b) \geq v(MIP)$. to this end, put

$$(31) \qquad F(v) = \inf \left\{ (c-\lambda A)\, x + (d - \lambda B)y \;\middle|\; \begin{array}{l} Ax + By = v, x,y \geq 0 \\ \text{and } x \text{ integer} \end{array} \right\}$$

F is subadditive since it is a value function (see [6]) and F is non-negative since $c - \lambda A \geq 0$ and $d - \lambda B \geq 0$.

Let (x^o,y^o) be a feasible solution to (MIP) with $F(b) = (c - \lambda A)\, x^o + (d - \lambda B)y^o$ ((x^o,y^o) exists by [9]). We have $v(MIP) \leq cx^o + dy^o$ and also

$$(32) \qquad cx^o + dy^o = (c-\lambda A)x^o + (d-\lambda B)y^o + \lambda Ax^o + \lambda By^o$$
$$= F(b) + \lambda b .$$

Set $H(v) = \min \{ F(v), F(b)\}$. Then $v(MIP) \leq H(b) + \lambda b$, as $H(b) = F(b)$. By Lemma 1.8, $H\epsilon\Sigma$. We need only prove that H is a feasible solution to (D_λ), and we are done.

However, $H(v) \le F(v)$ for all v. Hence $H(a^{(j)}) \le F(a^{(j)}) \le c_j - \lambda a^{(j)}$ for all j, and $\bar{H}(b^{(k)}) \le \bar{F}(b^{(k)}) \le d_k - \lambda b^{(k)}$ for all k, the latter by (25). Indeed, H is feasible in (D_λ) .

Q.E.D.

As a consequence of Theorem 2.4, if we define a dual program (D) to be (D_λ) with λ treated as a variable, we again have equality of values $v(D) = v(MIP)$. Indeed, for any feasible solution λ, F to D, since $F \epsilon \Sigma$ it is easy to show that λ is a solution to (DLP), and Theorem 2.4 applies. The proof of Theorem 2.4 shows that, by chosing $\lambda = \bar{\lambda}$, with $\bar{\lambda}$ an optimum in (DLP), we can take F in $(D_{\bar{\lambda}})$ to have an upper bound equal to the "integrality gap" $G(b) - L(b)$. This is the smallest bound possible on F for any of the dual programs (D_λ).

Clearly, information on the <u>gap function</u> $GP(b) = G(b) - L(b)$ is of value. By Corollary 1.3. it is a bounded, non-negative function.

<u>Example</u>: Consider the pure-integer program

$$\text{(33)} \qquad \begin{aligned} &\text{minimize } -(x_1 + x_2 + x_3) \\ &\text{subject to } \begin{pmatrix} 1 \\ 0 \end{pmatrix} x_1 + \begin{pmatrix} 0 \\ 1 \end{pmatrix} x_2 + \begin{pmatrix} -2 \\ 4 \end{pmatrix} x_3 = b \\ &\qquad x_1, x_2, x_3 \ge 0 \text{ and integer} \end{aligned}$$

for the three settings of b given by

$$\text{(34)} \qquad b^{(1)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad b^{(2)} = \begin{pmatrix} -2 \\ 5 \end{pmatrix}, \quad b^{(3)} = \begin{pmatrix} -1 \\ 6 \end{pmatrix}$$

For $b = b^{(1)}$, an optimum to (33) is (1,1,0), and this solution is also the optimum to (LP); hence $GP(b^{(1)}) = G(b^{(1)}) - L(b^{(1)}) = 0$. For $b = b^{(2)}$, an optimum to (33) and to (LP) is (0,1,1), and again $GP(b^{(2)}) = 0$. However, for $b = b^{(3)}$ an optimum to (LP) is (0,4,1/2), and an optimum to (33) is (1,2,1). Therefore, $GP(b^{(3)}) = -4 + 4\ 1/2 = 1/2$, and the subadditivity relation

(35) $$GP(b^{(3)}) \leq GP(b^{(1)}) + GP(b^{(2)}) = 0$$

clearly fails, even though $b^{(3)} = b^{(1)} + b^{(2)}$.

Despite the lack of global subadditivity, GP is subadditive in certain regions that we now define. Put

(36) $$S_{\bar{\lambda}} = \{b \epsilon S \mid \bar{\lambda} \text{ is optimal in the dual program (DLP) to (LP)}\}$$

For any specific $b^o \epsilon S$, there is some $\bar{\lambda}$ optimal in (DLP). By Lemma 1.1, $\bar{\lambda}$ is still optimal in (DLP) for all $b \epsilon S$ such that b is a non-negative combination of the columns which have positive variables in an optimal solution to (LP) with r.h.s. b^o. In specific, $\bar{\lambda}$ is still optimal in (DLP) for any r.h.s. b for which the optimal linear programming basis for r.h.s. b^o is still feasible, and hence if such a $b \epsilon S$ then also $b \epsilon S_{\bar{\lambda}}$.

The preceeding remarks are intended to give an intuitive idea of the nature and extent of $S_{\bar{\lambda}}$. Roughly speaking, $S_{\bar{\lambda}}$ includes all those feasible r.h.s. b which have some common optimal linear programming basis which is non-degenerate for some r.h.s. Clearly, if $S_{\bar{\lambda}} \neq \emptyset$ then $0 \epsilon S_{\bar{\lambda}}$ and S_{λ} is closed under addition, just as is true of S. However, $S_{\bar{\lambda}}$ can be more extensive than the region in which one linear programming basis is optimal. For instance, if there are dual non-degenerate linear programming basic optima for some r.h.s. b, the corresponding $\bar{\lambda}$ has $S_{\bar{\lambda}}$ include both regions in which either basis is optimal.

More information on the sets $S_{\bar{\lambda}}$, as well as the important global question of how they "fit together," is given in the next result. In what follows, dim(T) is the dimension of a set T; faces of polyhedra are are defined in [10], [13].

Theorem 2.5: Let α, β be optima in the dual program (DLP) for (possibly distinct) r.h.s. b.

(1) $\text{conv}(S_\alpha)$ is a polyhedral cone, and $\text{conv}(S_\alpha) \cap S = S_\alpha$

(2) $\text{conv}(S_\alpha) \cap \text{conv}(S_\beta)$ is a face (possibly empty) of both $\text{conv}(S_\alpha)$ and $\text{conv}(S_\beta)$.

Proof: By the Finite Basis Theorem (see [10], [13]), the dual polyhedron has a representation of the form

(37) $$\{ \lambda | \lambda A \le c, \lambda B \le d\} = \text{conv}\{v^{(1)},\dots,v^{(\theta)}\} + \text{cone}\{w^{(1)},\dots,w^{(\sigma)}\}$$

We may assume $\theta > 0$. Hence the linear form λb is bounded above subject to $\lambda A \le c$, $\lambda B \le d$ if and only if

(38) $$w^{(j)}b \le 0, \; j = 1,\dots,\sigma$$

(these inequalities being vacuous if $\sigma = 0$). When (38) holds, the maximum of λ b subject to these constraints is $\max \{v^{(i)}b \mid i = 1,\dots,\theta\}$.

Therefore, α is an optimum in the dual program if and only if both (38) holds and also

$$\alpha b \geq \beta b$$

$$(39)_\alpha \qquad \alpha b \geq v^{(i)} b \quad , \quad i = 1,\dots,\theta$$

(The inequality $\alpha b \geq \beta b$ is actually redundant here).

By the integrality hypothesis (SA), all the vectors $v^{(i)}, w^{(j)}$ in (37) - $(39)_\alpha$ can be assumed integral.

We now claim that $\mathrm{conv}(S_\alpha)$ is precisely the polyhedral cone C_α of b described by (38), $(39)_\alpha$.

Since the defining inequalities (38), $(39)_\alpha$ of C_α are in integral quantities, $C_\alpha = \mathrm{conv}\ \{\ b \in C_\alpha \mid b \text{ integral}\ \}$. However, if $b \in C_\alpha$ and b is integral, (LP) has a rational optimum with rationals having denominator $D \geq 1$ for an integral D, hence $Db \in S$ by Lemma 1.1. Since C_α is a cone, we obtain $C_\alpha = \mathrm{conv}\{\ b \in C_\alpha \mid b \in S\ \} = \mathrm{conv}(S_\alpha)$, as we claimed.

From this claim, if $b \in \mathrm{conv}(S_\alpha) \cap S$, then $b \in C_\alpha$ and $b \in S$, so indeed $b \in S_\alpha$.

Next, the claim shows that $\mathrm{conv}(S_\alpha) \cap \mathrm{conv}(S_\beta)$ is defined by (38), $(39)_\alpha$ plus the equality

$$(40) \qquad \alpha b = \beta b$$

Now the equality (40) occurs as an inequality of $(39)_\alpha$, hence requiring (40) will give a face (possibly empty) of $C_\alpha = \mathrm{conv}(S_\alpha)$.

Q.E.D.

It follows from Theorem 2.5 that the polyhedral cone C of all b satisfying (38) (i.e., those b for which (LP) is consistent) can be represented as follows. If d = dim(C), let $\alpha(1),\ldots,\alpha(\rho)$ be distinct points among $\{v^{(1)},\ldots,v^{(\delta)}\}$ of (37) for which $\dim(\mathrm{conv}(S_{\alpha(i)})) = d$, such that $i \neq j$ implies $S_{\alpha(i)} \neq S_{\alpha(j)}$. Then we have

$$(41) \qquad C = \bigcup_{i=1}^{\rho} \mathrm{conv}(S_{\alpha(i)})$$

and if $i \neq j$, then $\mathrm{conv}(S_{\alpha(i)}) \cap \mathrm{conv}(S_{\alpha(j)})$ is a face of $\mathrm{conv}(S_{\alpha(i)})$ with dimension not exceeding $(d - 1)$.

<u>Theorem 2.6</u>:

If $\bar{\lambda}$ is an optimum to (DLP), then GP is subadditive on $S_{\bar{\lambda}}$.

In fact, for $b \in S_{\bar{\lambda}}$ we have

$$(42) \qquad GP(b) = \min \left\{ (c - \bar{\lambda}A)x + (d - \bar{\lambda}B)y \;\middle|\; \begin{array}{l} Ax + By = b,\ x,\ y \geq 0 \\ \text{and } x \text{ integer} \end{array} \right\}$$

<u>Proof</u>: Exactly as in (31), (32), we see that if the minimum on the r.h.s. of (42) is attained at (x^o, y^o), we have

$$(43) \qquad G(b) = cx^o + dy^o = (c - \bar{\lambda}A)x^o + (d - \bar{\lambda}B)y^o + \bar{\lambda}b \,.$$

For $b \in S_{\bar{\lambda}}$, we have $L(b) = \bar{\lambda}b$. Using this and (43), we obtain $GP(b) = (c - \bar{\lambda}A)x^o + (d - \bar{\lambda}B)y^o$, which is (42), and we recall [6] that value functions are subadditive.

Q.E.D.

We conclude with an alternate representation of the gap function GP as the finite limit of value iteration applied to any "initial" non-negative subadditive function $F_o = F$ which does not exceed GP, and show that, at each step of value iteration, the current subadditive function F_n (see below) is "improved" (see Theorem 2.8). The specific value iteration here is the one corresponding to a shortest-path problem in a certain graph, as formulated in a dynamic programming context; we next describe the graph, which is closely related to Gomory's "round-off-problem" [4] and involves enumeration on the non-basic variables of an ℓ.p. tableau.

Let the linear programming relaxation of (IP) be solved as described previously, and let $A = [\, U \vdots V \,]$ be partitioned into basic and non-basic columns. Note that (IP) is equivalent to

$$\text{(IP)}' \qquad \begin{array}{ll} \min & c_U\bar{b} + (\, c_V - c_U\,\bar{A}\,)\, x_V \\ \text{subject to} & x_U + \bar{A}x_V = \bar{b} \\ & x_U\,,\, x_V \geq 0 \text{ and integer} \end{array}$$

where $\bar{A} = U^{-1}V$ and $\bar{b} = U^{-1}b$. By dual optimality, we have $\gamma = c_V - c_U\bar{A} \geq 0$. For notational purposes, we write the columns of $\bar{A}$ as follows:

$$\bar{A} = \left[\bar{a}^{(1)}, \bar{a}^{(2)}, \ldots, \bar{a}^{(t)}\right] \tag{44}$$

where $t = r - m \geq 0$. Clearly, $\bar{b} \geq 0$. Note that $L(b) = c_U B^{-1} b = c_U\bar{b}$. In what follows, we assume $b \epsilon S$, i.e., (IP) is consistent.

The node set P of the graph G = (P,E) consists of all vectors $v \in R^m$ having a representation

$$v = u + \sum_{j=1}^{t} p_j \bar{a}^{(j)} \tag{45}$$

where the p_j are integers and u is a vector of integers. As regards the edge set E of G = (P,E), a *directed arc* (v,w) exists for $v,w \in P$ if and only if either

$$v = \bar{a}^{(j)} + w \text{ for some } j = 1,\ldots,t \tag{46a}$$

or

$$v = e_k + w \text{ for some unit vector } e_k,\ k = 1,\ldots,m. \tag{46b}$$

In (46a), the length of this arc is γ_j, the j-th component of γ . In (46b), the length of this arc is zero. (If multiple arcs exist from v to w, these can be replaced by one possessing minimum length). Clearly, the set P is a group under addition and $\bar{b} \in P$.

In addition to nodes and edges, there is a distinguished subset $T \subseteq P$ of the set of nodes P, consisting of those $u \in P$ which are vectors of non-negative integers. Clearly, $T \neq \emptyset$, $\bar{b} \in P$, and there is a path from $\bar{b} \in P$ to T, since if (x_U^*, x_V^*) is a feasible solution to (IP), then $\bar{b} = x_U^* + Ax_V^*$ shows $\bar{b} \in P$ and $x_U^* = \bar{b} - \bar{A}x_V^*$ describes a path from $\bar{b}$ to $x_U^* \in T$. More generally, there is a path from $U^{-1}v \in P$ to T if and only if (IP) is consistent with r.h.s. v.

Let LH(v) be the length of a shortest path from v to T, for $v \in P$, where we set LH(v) = + ∞ if no path exists from $v \in P$ to T.

Let $\bar{\lambda}$ be the optimum in the linear dual to (IP)' . The next result is elementary, and we omit its proof.

<u>Proposition 2.7:</u> If $v \in S_{\bar{\lambda}}$ then $G(v) = \bar{\lambda}v + LH(U^{-1}v)$ and

(47) $$LH(U^{-1}v) = GP(v)$$

The following is a value-iteration scheme which begins at an "initial function" $F_o = F$ defined on P. Given F_n, we obtain F_{n+1} by:

(48) $$F_{n+1}(v) = \begin{cases} 0, & \text{if } v \in T \\ \min \{ \gamma_j + F_n(v - \bar{a}^{(j)}) \} , & \text{if } v \notin T \end{cases}$$

In what follows, for two functions F and G, we shall write $F \leq G$ to abbreviate $F(v) \leq G(v)$ for all $v \in P$.

Recall that a function F is <u>subadditive</u> on P if

(49) $$F(v + w) \leq F(v) + F(w) \qquad \text{for } v,w \in P.$$

See also [4], [6], [8].

<u>Theorem 2.8:</u>

If $F_0 = F$ is a subadditive function, and $0 \leq F \leq LH$ then every F_n defined by (48) is subadditive and $F_n \leq LH$. Also,

(50) $$F_o \leq F_1 \leq F_2 \leq \cdots$$

If, in addition, the vector γ is strictly positive, then for some finite N we have

$$F_N(v) = LH(v) \tag{51}$$

whenever $v \in S_{\bar{\lambda}}$.

<u>Proof</u>: We show $F_n \leq LH$ by induction on n. For n = 0, this is an hypothesis. To go from n to (n+1), if $v \in T$ then $F_{n+1}(v) = 0 = LH(v)$, while if $v \notin T$ then by (48) and $F_n \leq LH$

$$F_{n+1}(v) \leq \min_{j=1,\ldots,t} \{\gamma_j + LH(v - \bar{a}^{(j)}) \} = LH(v) \tag{52}$$

F_n is non-negative for n = 0 by hypothesis. To go from n to (n+1), if $v \in T$ then $F_{n+1}(v) = 0$ while by (48) for $v \notin T$ we have by induction

$$F_{n+1}(v) = \gamma_k + F_n(v - \bar{a}^{(j)}) \geq 0 + 0 = 0 \tag{53}$$

for at least one index k = 1,...,t . Hence, all F_n are non-negative.

F_n is subadditive for n = 0 by hypothesis. To go from n to (n+1), we first use the hypothesized subadditivity of F_n to prove

$$F_{n+1}(v) \geq F_n(v) \tag{54}$$

If $v \in T$, then $0 = LH(v) \geq F_n(v) \geq 0$ and $F_{n+1}(v) = 0$ by (48), hence (54) holds. If $v \notin T$, then since $v = \bar{a}^{(j)} + (v - \bar{a}^{(j)})$ for j = 1,...,t, subadditivity of F_n gives

(55) $$F_n(v) \leq F_n(\bar{a}^{(j)}) + F_n(v - \bar{a}^{(j)})$$
$$\leq \gamma_j + F_n(v - \bar{a}^{(j)}) \qquad , j = 1,\ldots,t$$

using $F_n(\bar{a}^{(j)}) \leq LH(\bar{a}^{(j)}) \leq \gamma_j$, $j = 1,\ldots,t$. Now

(48) and (55) give (54).

Using (54), and supposing F_n subadditive, if $v + w \in T$ then $F_{n+1}(v + w) = 0 \leq F_{n+1}(v) + F_{n+1}(w)$ since F_{n+1} is non-negative. However, if $v+w \notin T$, then by (48) we have

(56) $$F_{n+1}(v + w) \leq \gamma_k + F_n(v + w - \bar{a}^{(k)})$$

$$\leq \gamma_k + F_n(v - \bar{a}^k) + F_n(w)$$

$$= F_{n+1}(v) + F_n(w) \leq F_{n+1}(v) + F_{n+1}(w) .$$

In (56), we have assumed $v \notin T$ WLOG (for $v, w \in T$ implies $v + w \in T$) and have chosen k so that $F_{n+1}(v) = \gamma_k + F_n(v - \bar{a}^{(k)})$ by (48). Clearly, (56) shows that F_{n+1} is also subadditive and completes our inductive proof that all F_n are subadditive. Then (50) follows since (54) is implied by the subadditivity of F_n.

Let us now assume, in addition, that γ is strictly positive, and let $\epsilon > 0$ be chosen so that $\gamma_j \geq \epsilon$ for $j = 1,\ldots,t$. We now prove by induction on n that

(57) $$F_{n+1}(v) \geq \min \{ n\epsilon, LH(v) \} .$$

For n = 0, (69) follows since $F_o = F \geq 0$. To go from n to (n+1), if $v \epsilon T$ then $F_{n+1}(v) = 0$ and $LH(v) = 0$. If $v \notin T$, then by (48)

(58) $$F_{n+1}(v) = \gamma_k + F_n(v - \bar{a}^{(k)})$$

for some k = 1,...,t. In the event that $F_n(v - \bar{a}^{(k)}) \geq (n-1)\epsilon$, we have from (58) and $\gamma_k \geq \epsilon$ that $F_{n+1}(v) \geq \min \{ n\epsilon, LH(v) \}$. On the other hand, if $F_n(v - \bar{a}^{(k)}) \geq LH(v - \bar{a}^{(k)})$, we have from (58)

(59) $$F_{n+1}(v) \geq \min_{j=1,\ldots t} \{\gamma_j + LH(v - \bar{a}^{(j)})\} = LH(v)$$

$$\geq \min \{n\epsilon, LH(v)\} .$$

This completes our proof of (57).

Let N be chosen to that $N\epsilon \geq \Delta$, where Δ is the ℓ.h.s. in (9) of Corollary 1.3. By Proposition 2.7 , if v is a feasible r.h.s. then $LH(v) = GP(v) \leq N\epsilon$, so $\min \{N\epsilon, LH(v)\} = LH(v)$. As $F_N \leq LH$, (57) gives $F_N(v) = LH(v)$ for v feasible, and (51) holds.

Q.E.D.

December 10, 1976

References

[1] C.E. Blair, "Extensions of Subadditive Functions Used in Cutting-Plane Theory," MSRR no. 360, Carnegie-Mellon University, December 1974.

[2] C.E. Blair and R.G. Jeroslow, "The Value Function of a Mixed Integer Program: I," MSRR no. 377, Carnegie-Mellon University, October 1975.

[3] E.W. Dijkstra, "A Note on Two Problems in Connexion with Graphs," Numerische Mathematik I (1959), pp. 269-271.

[4] R.E. Gomory, "Some Polyhedra Related to Combinatorial Problems," Linear Algebra and Its Applications 2 (1969), pp. 451-558.

[5] R.G. Jeroslow, "The Principles of Cutting-Plane Theory: Part I," Carnegie-Mellon University, February 1974.

[6] R.G. Jeroslow, "Minimal Inequalities," MSRR no. 362, Carnegie-Mellon University, April 1975.

[7] R.G. Jeroslow, "Solving Nearby Integer Programs," MSRR no 392, Carnegie-Mellon University, April 1976.

[8] E.L. Johnson, "The Group Problem for Mixed Integer Programming," Mathematical Programming Studies, Study 2, December 1974, pp. 137-179.

[9] R.R. Meyer, "On the Existence of Optimal Solutions to Integer and Mixed-Integer Programming Problems," Mathematical Programming 7 (1974), pp. 223-235.

[10] R.T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, N.J., 1970. 432+ pages.

[11] R.T. Rockafellar, "Augmented Lagrange Multiplier Functions and Duality in Nonconvex Programming," SIAM Journal on Control 12 (1974), pp. 268-285.

[12] R.T. Rockafellar, "A Dual Approach to Solving Nonlinear Programming Problems by Unconstrained Optimization," Mathematical Programming 5 (1973), pp. 354-373.

[13] J. Stoer and C. Witzgall, Convexity and Optimization in Finite Dimensions: I, Springer-Verlag, New York, 1970. 268+ pages.

[14] L.A. Wolsey, "Extensions of the Group Theoretic Approach in Integer Programming," Management Science 18 (1971), pp. 74-83.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER M.S.R.R. 397	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER (14) RR-397
4. TITLE (and Subtitle) The Value Function of a Mixed Integer Program: II (6)		5. TYPE OF REPORT & PERIOD COVERED December, 1976
		6. PERFORMING ORG. REPORT NUMBER M.S.R.R. 397
7. AUTHOR(s) (10) C.E. Blair and R.G. Jeroslow		8. CONTRACT OR GRANT NUMBER(s) (15) N00014-75-C-0621, NSF-GP-37510-X1
9. PERFORMING ORGANIZATION NAME AND ADDRESS GSIA Carnegie-Mellon University, Pittsburgh, PA		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 047-048
11. CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Programs Office of Naval Research (Code 458) Arlington, Virginia 22217		12. REPORT DATE (11) December, 1976
		13. NUMBER OF PAGES 25 (12) 29p.
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (9) Research rept.,		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Integer programming, Cutting-planes, Convex analysis, Duality, Subadditivity

It is proved

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

~~We prove~~ that the gap in optimal value, between a mixed-integer program in rationals and its corresponding linear programming relaxation, is bounded as the right-hand-side is varied. In addition, a variant of value iteration is shown to construct subadditive functions which resolve a pure-integer program when no dual degeneracy occurs. These subadditive functions provide solutions to subadditive dual programs for integer programs which are given here, and for which the values of primal and dual problems are equal.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

403426