Reprinted from the Proceedings of the National Academy of Sciences
Vol. 53, No. 2, pp. 260-265. February, 1965.

ON THE RELATION BETWEEN INTEGER AND NONINTEGER SOLUTIONS TO LINEAR PROGRAMS*

BY R. E. GOMORY

THOMAS J. WATSON RESEARCH CENTER, YORKTOWN HEIGHTS, NEW YORK

Communicated by R. Courant, December 22, 1964

We will refer to the ordinary linear programming problem

$$\begin{aligned} \text{maximize} \quad z_1 &= cx \\ Ax &= b, x \geq 0 \end{aligned} \tag{1}$$

as problem $P1$. In (1) b is an integer m-vector, c is an $m + n$ vector, and A is an $m \times (m + n)$ integer matrix. x is an $m + n$ vector, all of whose components are required to be nonnegative. We assume that A is of the form (A', I) with I an $m \times m$ identity matrix, so that in (1) $Ax = b$ is equivalent to the m inequalities in n variables $A'x' \leq b$. We will say that x is feasible if it satisfies the equality and nonnegativity conditions of (1) and optimal if it also maximizes.

A problem closely related to $P1$ is the integer programming problem $P2$ which is $P1$ with the added condition that the components of x be integers. Because of the comparative ease with which $P1$ is solved[1] and the comparative difficulty of $P2$,[2, 3] it is natural to consider getting from the solution of $P1$ to the solution to $P2$ by some sort of a "rounding" process through which the noninteger components of the x solving $P1$ are rounded either up or down to produce a solution to $P2$. This procedure seems particularly plausible when the components x_i of x are reasonably large numbers. However, it is easily shown by examples that a nearest-neighbor rounding process cannot generally produce the optimal solution to $P2$. These examples are neither pathological nor uncommon; it is simply not the case that the optimal solution can be obtained by simple rounding to some vector x' with $|x'_i - x_i| < 1$, even if the rounding is followed by some sort of optimization on the residual problem and even if the b and x of (1) become arbitrarily large.

Nevertheless, there is a close connection between the optimal solutions to $P1$ and $P2$ for a wide range of right-hand sides b. We first give some theorems on this connection and then an algorithm which for these b obtains the optimal solution of $P2$ from the optimal solution to $P1$.

If B is a basis, i.e., an $m \times m$ nonsingular submatrix of A, we will assume that A has been rearranged and partitioned into matrices B and N with $A = (B,N)$. We will also partition $x = (x_B, x_N)$ and $c = (c_B, c_N)$. The columns of A will be referred to as α_i, $B = (\alpha_1, \ldots, \alpha_m)$. We confine ourselves to right-hand side vectors b in that part of m-space for which (1) is solvable. If B is the optimal basis for $P1$ with right-hand side b, then it is also the optimal basis for all b' such that $B^{-1}b' \geq 0$. These b form a cone in m-space, and in fact all solvable m-space is partitioned into such cones K^B. On removing from K^B all points within a distance d of its boundary, we have the reduced cone $K^B(d)$. With this notation we can now state Theorem 1.

THEOREM 1. *Let* $l = \max |\alpha_i|$, $i = m + 1, \ldots, m + n$, $D = |\det B|$, *and* $z_1(b)$ *be the value of the solution to* $P1$. *Then if* $b \epsilon K^B(l(D - 1))$, *the value* $z_2(b)$ *of the solution to* $P2$ *is given by*

$$z_2(b) = z_1(b) + \varphi^B(b), \tag{2}$$

and an optimal solution vector is given by

$$x(b) = (x_B(b), x_N(b)) = (B^{-1}(b - Ny^B(b)), y^B(b)), \tag{3}$$

where both the scalar function $\varphi^B(b)$ *and the n-vector function* $y^B(b)$ *are m-periodic, i.e.,* $\varphi^B(b + \alpha_i) = \varphi^B(b)$, $i = 1, \ldots, m$, *and* $y^B(b + \alpha_i) = y^B(b)$, $i = 1, \ldots, m$.

The periodicity means that the values of $\varphi^B(b)$ and $y^B(b)$ depend only on the position of b relative to the lattice $\mathfrak{L}_B$ of points generated by integer combinations of $\alpha_1, \ldots, \alpha_m$. This is equivalent to saying that φ^B is a function on the factor module $M(I)/M(B)$ where $M(I)$ is the module of all integer points in m-space and $M(B)$ the module of integer combinations of the α_i, $i = 1, \ldots, m$.

Although (2) and (3) have just the form one would expect if rounding were possible, the integer solution $x(b)$ is generally not a continuation of a rounded nearest-neighbor solution. It is instead a continuation from a point $p = b - Ny^B(b)$ which is on $\mathfrak{L}_B$. A measure of the distance from p to b is given by Theorem 2.

THEOREM 2. *If* $b \epsilon K^B(l(D - 1))$, *then the optimal solution vector* $x(b)$ *has the property*

$$\sum_{i=m+1}^{i=m+n} x_i = \sum_{i=1}^{i=n} y_i \leq D - 1.$$

We next discuss the arithmetic work involved in actually obtaining $\varphi^B(b)$ and

$y^B(b)$. The calculation may be broken into two parts, of which the first is a standard calculation.

The factor module $M(I)/M(B)$ is a finite additive group having D elements. By the methods of references 4 or 5 we calculate $M(I)/M(B)$ as the direct sum of cyclic groups of known orders. The arithmetic work involved is bounded by $2m(m^2 + 2m)\log_2 D$. Or, alternatively, given B^{-1} as a starting point, the standard form of $M(I)/M(B)$ can be obtained in at most $2r(m^2 + 2m)\log_2 D$ arithmetic steps, where $r \leq m$ is the rank of $M(I)/M(B)$. If the factor group is cyclic, r is 1.

This calculation also provides explicitly a means of mapping integer m-vectors onto corresponding elements of $M(I)/M(B)$. If we call this mapping f, then $\bar{p} = fp$ can be obtained for any integer m-vector p by at most $m^2 + 2m$ arithmetic steps.

Since both $\varphi^B(b)$ and $y^B(b)$ are m-periodic, one can obtain, by periodicity, values $\varphi^B(b)$ and $y^B(b)$ for all b if they are known for one period.

THEOREM 3. *There are D distinct b values in one period. If $M(I)/M(B)$ has been put in standard form and the group elements $\bar{\alpha}_i = f\alpha_i$ obtained for $i = m + 1, \ldots, m + n$, then the values of $\varphi^B(b)$ can be computed for all b in one period in less than $7nD$ elementary arithmetic steps. The values of $y^B(b)$ for a particular b can be computed in n more steps or the values for all b in one period in nD more steps.*

The arithmetic steps referred to here are operations such as the addition and subtraction of real numbers, comparison of real numbers, or the addition and subtraction of elements of $M(I)/M(B)$. We now turn to the proofs of these theorems.

In reference 2 it was pointed out that $M(I)/M(B)$ is isomorphic to the group F generated by the rows of the matrix $B^{-1}A$ with the entries being replaced by the corresponding entries modulo 1. These "fractional rows" then provided the basic inequalities for the methods of reference 2 and, in a less evident way, for reference 3. As was remarked in reference 2, similar reasoning shows that $M(I)/M(B)$ is also isomorphic to the group generated by the columns of $B^{-1}A$ with coefficients being treated the same way, i.e., reduced to proper fractions. As $B^{-1}A$ is the simplex tableau provided by the simplex method in solving (1), each column has associated with it a relative cost factor. This representation suggests the following problem involving maximization over $M(I)/M(B)$.

$$\max \sum_{i=1}^{i=n} c^*_{i+m} y_i$$

$$\sum_{i=1}^{i=n} \bar{\alpha}_{i+m} y_i = \bar{b},\ y_i \geq 0 \text{ and integer.} \tag{4}$$

Here $\bar{\alpha}_i$ and $\bar{b}$ are the elements of $M(I)/M(B)$ corresponding to the vectors α_i and b, and c^*_i is the relative cost $c^*_i = c_i - c_B B^{-1}\alpha_i$.

It is a fundamental property of linear programming that all c^*_i associated with an optimal basis are ≤ 0; so the maximum in (4) does exist for optimal bases (though not for other bases).

Since the $\bar{\alpha}_i \epsilon M(I)/M(B)$, a group with D elements, $D\bar{\alpha}_i = \bar{o}$; so for a minimal solution to (4) it is only necessary to consider y_i satisfying $0 \leq y_i < D$. We will indicate later how (4) can, in fact, be solved for all $\bar{b}$ in a total of $7nD$ elementary steps.

If b is an integer m-vector, define $\varphi(b)$ as the value of the solution of (4) with $\bar{b} = fb$. This is the φ^B of Theorem 1.

One of the properties of φ is immediate. Clearly, $\varphi(b + \alpha_i) = \varphi(b), i = 1, \ldots, m$. Also $\varphi(b) + c_BB^{-1}b \geq z_2(b)$. For if (x'_B, x'_N) is a feasible integer solution to (1), then the corresponding cost $c(x'_B, x'_N) = c_Bx'_B + c_Nx'_N$ can be expressed in terms of the x'_N only by using the relation

$$Bx'_B + Nx'_N = b, \tag{5}$$

which yields

$$c(x'_B, x'_N) = c_BB^{-1}b + (c_N - c_BB^{-1}N)x'_N.$$

${}_BB^{-1}b$ is $z_1(b)$, so

$$c(x'_B, x'_N) = z_1(b) + \sum_{i=m+1}^{i=m+n} c^*_i x'_i. \tag{6}$$

Applying the homomorphism f to (5), B disappears and we get

$$\sum_{i=m+1}^{i=m+n} \bar{\alpha}_i x'_i = \bar{b},$$

so x'_N is a feasible solution to (4). Hence

$$\sum_{i=m+1}^{i=m+n} c^*_i x'_i \leq \varphi(b);$$

so from (6), $c(x'_B, x'_N) \leq z_1(b) + \varphi(b)$, and since this holds for all such (x'_B, x'_N),

$$z_2(b) \leq z_1(b) + \varphi(b). \tag{7}$$

Now let us consider the y that solves (4). For this y

$$\sum_{i=1}^{i=n} c^*_{i+m} y_i = \varphi(b).$$

We extend this to a solution to (5) by choosing $x_B = B^{-1}(b - Ny)$. Since y solves (4), $(b - Ny)\epsilon\mathcal{L}_B$, so x_B will be integral. If x_B is also nonnegative, (x_B, y) solves (1) and, in fact, is the optimal integer solution since its cost, by (6), is $z_1(b) + \varphi(b)$, which, by (7), establishes optimality.

To establish conditions for the nonnegativity of x_B we need the following lemma.

LEMMA. *There is an optimal solution to (4) with*

$$\sum_{i=1}^{i=n} y_i \leq D - 1.$$

Proof: If the y_i are the components of that optimal solution having $\sum y_i$ minimal, form any sequence of the $\bar{\alpha}_{i+m}$ in which each $\bar{\alpha}_{i+m}$ appears exactly y_i times. Then form the partial sum $\bar{S}_p$ of the first p elements of the sequence. We include $\bar{S}_0 = \bar{o}$. If the sequence has more than $D - 1$ elements, there are more than D $\bar{S}_p$; so there must be a p and p', $p < p'$ for which $\bar{S}_p = \bar{S}_{p'}$. The elements, between p and p' in the sequence total $\bar{o}$ and can be deleted. The remaining elements form a new solution which contradicts either optimality or minimal total $\sum y_i$.

A related argument can be used to prove the following which, though not used in the proofs of Theorems 1, 2, or 3, bears on the multiplicity of solutions to integer programs.

THEOREM 4. *If* $\prod_{i=1}^{i=n}(y_i + 1) > D - 1$, *then the solution to (4) is not unique.*

Returning to the proof, it follows from the lemma that $\|Ny\| \leq (D-1)l$. Hence, if $b \epsilon K^B(l(D-1))$, $b - Ny$ will be in K^B and so x_B will be nonnegative. This establishes Theorem 1.

Theorem 2 now follows at once from the lemma.

To establish Theorem 3, we now turn to the actual computation of y. Guided by dynamic programming,[6] we define $\varphi_s(\bar{p})$, $\bar{p} \epsilon M(I)/M(B)$ as the solution to

$$\max \sum_{i=1}^{i=s} c^*_{i+m} y_i$$

$$\sum_{i=1}^{i=s} \bar{\alpha}_{i+m} y_i = \bar{p}.$$

We have recursively

$$\varphi_s(\bar{p}) = \max \{\varphi_s(\bar{p} - \bar{\alpha}_{s+m}) + c^*_{s+m}, \varphi_{s-1}(\bar{p})\}. \quad (8)$$

Let us assume that $\varphi_{s-1}(\bar{p})$ is known for all $\bar{p} \epsilon M(I)/M(B)$. Then we can compute recursively, starting with $\varphi_s(\bar{o}) = 0$,

$$\varphi(\bar{\alpha}_{s+m}) = \max \{\varphi_s(o) + c^*_{s+m}, \varphi_{s-1}(\bar{\alpha}_{s+m})\}$$

$$\varphi(r\bar{\alpha}_{s+m}) = \max \{\varphi_s(r\bar{\alpha}_{s+m} - \bar{\alpha}_{s+m}) + c^*_{s+m}, \varphi_{s-1}(r\bar{\alpha}_{s+m})\}, r = 1, 2, \ldots, D-1.$$

If $\bar{\alpha}_s$ is of order D, we will obtain all values $\varphi_s(\bar{p})$. If $\bar{\alpha}_s$ is of some order d which divides D, then $d\bar{\alpha}_s = \bar{o}$, and after d steps we return to $\bar{o}$. One then chooses some $\bar{p}$ not yet reached in the calculation (the standard form of the group is needed here), and setting $\varphi'_s(\bar{p}) = \varphi_{s-1}(\bar{p})$ computes $\varphi'_s(\bar{p} + r\bar{\alpha}_{s+m}) = \max \{\varphi'_s(\bar{p} + r\bar{\alpha}_{s+m} - \bar{\alpha}_{s+m}) + c^*_{s+m}, \varphi_s(p + r\bar{\alpha}_{s+m})\}$ $r = 1, \ldots, d$. After d steps, $d\bar{\alpha}_{s+m} = \bar{o}$, so we obtain a new value for $\varphi'_s(\bar{p})$ and then continue obtaining new values for $\varphi'_s(\bar{p} + r\bar{\alpha}_{s+m})$ the second time around. As soon as one of these new values agrees with the old, the calculation is stopped. It is not hard to show that: (*i*) the calculation will stop after q steps $d \leq q < 2d$; (*ii*) the $\varphi'_s(\bar{p} + r\bar{\alpha}_{s+m})$ values are the correct values $\varphi_s(p + r\bar{\alpha}_{s+m})$. This procedure is repeated for D/d starting points $\bar{p}$ to get values $\varphi_s(\bar{p})$ for all $p \epsilon M(I)/M(B)$.

If $M \leq \min_i c^*_i$, we can start with $\varphi_0(\bar{p}) = M$ for all $\bar{p}$. Then repeating the calculation leads to the calculation of $\varphi_n(\bar{p})$ for all $\bar{p}$ in at most $2nD$ elementary recursions each involving adding two group elements, looking up two values, adding two real numbers, and making one compare. To obtain the optimal solution with the smallest $\sum_i y_i$, one simply records with each $\varphi_s(\bar{p})$, when computed, the total $T_s(\bar{p}) = \sum_i y_i$ of the y_i of that solution. Clearly, $T_s(\bar{p}) = T_{s-1}(\bar{p})$, if the second term gives the maximum in (8) and $T_s(\bar{p}) = T_s(\bar{p} - \bar{\alpha}_{s+m}) + 1$, otherwise. In case of a tie in the maximum in (8), the term yielding the smaller $T_s(\bar{p})$ value should be chosen.

The solutions y_i are obtained by tracing back the recursion in the usual manner of dynamic programming. By proper recording, backtracking can be done even if the φ_{s-1} values are discarded, once the φ_s are known. These backtracking operations are virtually identical with those used in solving the knapsack problem,[7] and, in fact, work done with P. C. Gilmore on knapsack problems strongly suggested the results of this paper.

Finally, we note that the following steps, (*i*) solve $P1$ obtaining the opitmal B, (*ii*) put $M(I)/M(B)$ into standard form and identify the $\bar{\alpha}_i$, (*iii*) solve (4) obtaining y, (*iv*) compute $x_B = B^{-1}(b - Ny)$, will yield an optimal solution (x_B, y) if $x_B \geq 0$. It is not necessary for b to be in $K^B(l(D-1))$ to apply the procedure. The problems for which the procedure provides a solution are those for which those inequalities binding the solution of $P1$ alone determine the solution to $P2$.

* This work was supported in part by the Office of Naval Research under contract Nonr 3775-(00), NR 047040.

[1] Dantzig, George B., *Linear Programming and Extensions* (Princeton, N. J.: Princeton University Press, 1963).

[2] Gomory, R. E., "An algorithm for integer solutions to linear programs," *Recent Advances in Mathematical Programming*, ed. R. L. Graves and Philip Wolfe (McGraw-Hill, 1963), pp. 269–302.

[3] Gomory, R. E., "All-integer integer programming algorithm," *Industrial Scheduling*, ed. J. F. Muth and G. L. Thompson (Prentice-Hall, 1963), pp. 193–206.

[4] Zassenhaus, Hans, *The Theory of Groups* (New York: Chelsea Publishing Co., 1949).

[5] van der Waerden, B. L., *Modern Algebra* (New York: Frederick Ungar Publishing Co., 1950), vol. 2.

[6] Bellman, R., *Dynamic Programming* (Princeton, N. J.: Princeton University Press, 1957).

[7] Gilmore, P. C., and R. E. Gomory, "Multi-stage cutting stock problems of two and more dimensions," to appear in *Operations Res.*