THE NUMBER OF NECESSARY CONSTRAINTS IN AN INTEGER PROGRAM: A NEW PROOF OF SCARF'S THEOREM

by

Michael J. Todd

*This research was partially supported by National Science Foundation Grant ENG 76-08749 and by a CORE Fellowship at the Catholic University of Louvain, Belgium.
ABSTRACT

I give a new proof of Scarf's result that an integer program in $n$ variables has a set of binding constraints of cardinality at most $2^n - 1$. 
1. Introduction.

Consider the integer program

\[
\begin{align*}
\max & \quad a^0.h \\
\text{(P)} & \quad a^j.h \geq \beta_j, \quad j = 1, \ldots, m \\
& \quad h \in Z^n,
\end{align*}
\]

where \( Z^n \) is the set of \( n \)-dimensional integer vectors. We restrict the class of programs under consideration as follows.

Assumptions.

(a) (Integer Data) Each \( a^i \) lies in \( Z^n \) and each \( \beta_j \) is an integer.

(b) (Feasibility) Some \( h \in Z^n \) satisfies the constraints of (P).

(c) (Boundedness) The set \( \{ h \in \mathbb{R}^n | a^i.h \geq 0, \quad i = 0, \ldots, m \} \) consists of just the zero vector.

Requirements (b) and (c) ensure the existence of an optimal solution \( h^0 \) to (P). Assumption (a) is only included for simplicity.

Definition. The constraint \( a^j.h \geq \beta_j \) is necessary in (P) if there is some \( h \in Z^n \) satisfying all constraints of (P) except \( a^j.h \geq \beta_j \), with \( a^0.h > a^0.h^0 \). In other words, a constraint is necessary in (P) if its removal allows an increase in the optimal value of (P).

Our main result here is

Theorem 1. Under assumptions (a), (b) and (c), program (P) has at most \( Z^n - 1 \) necessary constraints.
The result above is not exactly in the form of Scarf's theorem [1]. To state Scarf's result we must define the concept of binding constraints.

**Definition.** The set \( J \subseteq \{1, \ldots, m\} \) of constraints of \((P)\) is a set of binding constraints if the optimal value of

\[
\max a^0 \cdot h \\
a^j \cdot h \geq \beta_j, \quad j \in J \\
h \in \mathbb{Z}^n
\]

is equal to that of \((P)\).

**Theorem 2.** (Scarf) Under assumptions (a), (b) and (c), program \((P)\) has a set of binding constraints of cardinality at most \(2^n - 1\).

The paper is organized as follows. In Section 2, we describe a transformation of one integer program into another and derive properties of the latter. In Section 3, we use this transformation repeatedly to prove Theorem 1. Section 4 contains the proof of Theorem 2, and indicates how assumption (a) can be removed. Section 5 compares the present proof with Scarf's.

2. An Integer Program Transformation

Suppose given an integer program

\[
\max c^0 \cdot h \\
(Q) \\
c^j \cdot h \geq \delta_j, \quad j = 1, \ldots, m \\
h \in \mathbb{Z}^n
\]
We assume that \( (Q) \) satisfies assumptions (a), (b) and (c) and that the constraint \( c^1.h \geq \delta_1 \) is necessary in \( (Q) \). Let \( \epsilon_0 \) be the optimal value of \( (Q) \). Then define \( (Q') \) as follows

\[
\text{max } c^1.h \\
(Q') \\
c^j.h \geq \delta_j, \quad j = 2, \ldots, m \\
c^0.h \geq \epsilon_0 + 1, \\
h \in \mathbb{Z}^n.
\]

**Lemma 1.** Under the assumptions above:

(i) \( (Q') \) satisfies assumptions (a), (b), (c);
(ii) The optimal value \( \epsilon_1 \) of \( (Q') \) is at most \( \delta_1 - 1 \);
(iii) The constraint \( c^0.h \geq \epsilon_0 + 1 \) is necessary in \( (Q') \);
(iv) For \( j = 2, \ldots, m \), if the constraint \( c^j.h \geq \delta_j \) is necessary in \( (Q) \), it is also necessary in \( (Q') \).

**Proof:** (i) The optimal value \( \epsilon_0 \) of \( (Q) \) is clearly integer; hence assumption (a) holds for \( (Q') \). Since \( c^1.h \geq \delta_1 \) was assumed necessary in \( (Q) \), there is some \( h \in \mathbb{Z}^n \) with \( c^j.h \geq \delta_j, \quad j = 2, \ldots, m \) and \( c^0.h \geq \epsilon_0 \); by integrality \( c^0.h \geq \epsilon_0 + 1 \) and \( (Q') \) is feasible.

Assumption (c) for \( (Q') \) is obvious.

(ii) First, (b) and (c) holding for \( (Q') \) ensure the existence of an optimal solution \( h^1 \) to \( (Q') \). If the optimal value of \( (Q') \) were at least \( \delta_1 \), \( h^1 \) would be feasible in \( (Q) \) with greater objective function value than \( \epsilon_0 \), a contradiction. Hence \( \epsilon_1 < \delta_1 \) and integrality gives \( \epsilon_1 \leq \delta_1 - 1 \).
(iii) If the constraint \( c^0 \cdot h \geq \epsilon_0 + 1 \) were removed from \((Q')\), the optimal solution to \((Q)\) would be feasible in \((Q')\) with value at least \( \delta_1 > \epsilon_1 \). Hence this constraint is necessary in \((Q')\).

(iv) Suppose the constraint \( c^2 \cdot h \geq \delta_2 \) is necessary in \((Q)\). Then there is a solution \( h \in \mathbb{Z}^n \) to \( c^1 \cdot h \geq \delta_1, \ c^j \cdot h \geq \delta_j, \ j = 3, \ldots, m \) and \( c^0 \cdot h > \epsilon_0 \); hence \( c^0 \cdot h > \epsilon_0 + 1 \). Thus the removal of \( c^2 \cdot h \geq \delta_2 \) from \((Q')\) allows the optimal value to increase at least to \( \delta_1 \) from \( \epsilon_1 \)---that is, \( c^2 \cdot h \geq \delta_2 \) is necessary in \((Q')\).

3. Proof of Theorem 1

Without loss in generality we may assume that the first \( q \) constraints in \((P)\) are necessary; we must show that \( q \leq 2^n - 1 \) under assumptions (a), (b) and (c). To do this we construct integer programs \((P_0), \ldots, (P_q)\) using the transformation of the last section.

\((P_0)\) is just the original problem:

\[
\begin{align*}
\max & \quad a^0 \cdot h \\
\text{subject to} & \quad a^j \cdot h \geq \beta_j, \quad j = 1, \ldots, m \\
& \quad h \in \mathbb{Z}^n.
\end{align*}
\]

We know \((P_0)\) has an optimal solution \( h^0 \) with value, say, \( a^0 \cdot h^0 = \gamma_0 \). In general, assume program \((P_i)\) with optimal solution \( h^i \) and optimal value \( \gamma_i \) has been constructed for \( 0 \leq i < j \). Then we define \((P_j)\) as follows:
\[
\begin{align*}
\max a^j.h \\
(P_j) & a^k.h \geq \beta_k, \quad k = j+1, \ldots, m \\
& a^i.h \geq \gamma_i+1, \quad i = 0,1,\ldots,j-1 \\
& h \in \mathbb{Z}^n.
\end{align*}
\]

Note that (P_j) is obtained from (P_{j-1}) as is (Q') from (Q) in Section 2. Thus using Lemma 1 inductively we have that (P_j) satisfies assumptions (a), (b) and (c); has an optimal solution \( h^j \) with value \( \gamma_j \leq \beta_j-1 \); and has as necessary constraints those indexed \( k = j+1, \ldots, q \) and \( i = 0, \ldots, j-1 \).

We now consider the optimal solutions \( h^0, \ldots, h^q \). If \( q \geq 2^n \) we have at least \( 2^{n+1} \) integer vectors. There are only \( 2^n \) elements in the \( n \)-dimensional vector space over the field of residues modulo 2. Hence two of the vectors, say \( h^i \) and \( h^j \), \( i < j \), agree modulo 2 in each component. Let \( h^* = (h^i+h^j)/2 \); then \( h^* \) is in \( \mathbb{Z}^n \).

Now \( h^i \) satisfies all constraints of (P_j) except \( a^i.h \geq \gamma_i+1 \); indeed \( a^i.h^i = \gamma_i \). Since \( a^i.h^j \geq \gamma_i+1 \) we have \( a^i.h^* \geq \gamma_i + \frac{1}{2} \), and by integrality, \( a^i.h^* \geq \gamma_i+1 \). Hence \( h^* \) satisfies all constraints of (P_j). Its value in this program is
\[
a^j.h^* = \frac{1}{2} a^j.i + \frac{1}{2} a^j.h^j \geq \frac{1}{2} \beta_j + \frac{1}{2} \gamma_j \geq \gamma_j + \frac{1}{2},
\]
since \( \beta_j \geq \gamma_j+1 \). Again, integrality implies \( a^j.h^* \geq \gamma_j+1 \). But this contradicts the fact that \( \gamma_j \) is the optimal value of (P_j). Hence \( q \leq 2^n-1 \) and the theorem is proved.

4. Proof of Theorem 2

Note first that Theorem 2 does not follow immediately from Theorem 1. If there are more than \( 2^n-1 \) constraints, Theorem 1 implies that one
may be removed. But the resulting program may no longer satisfy assumption (c) so that the reduction cannot continue. Instead of removing constraints, we therefore relax them.

Let $M$ be larger than $-\beta_j$ for $j = 1, \ldots, m$. If $m > 2^n - 1$, Theorem 1 implies that at least one constraint is unnecessary in (P), say the last. Hence the program

$$\begin{align*}
\max & \quad a^0 \cdot h \\
\text{s.t.} & \quad a^j \cdot h \geq \beta_j, \quad j = 1, \ldots, m-1 \\
& \quad a^m \cdot h \geq -M \\
& \quad h \in \mathbb{Z}^n
\end{align*}$$

(P')

has the same optimal value as (P). Also, the last constraint of (P') is clearly unnecessary. If $n-1 > 2^n - 1$ we may continue this process, obtaining finally

$$\begin{align*}
\max & \quad a^0 \cdot h \\
\text{s.t.} & \quad a^j \cdot h \geq \beta_j, \quad j \in J \\
& \quad a^j \cdot h \geq -M, \quad j \in \{1, \ldots, m\} \setminus J \\
& \quad h \in \mathbb{Z}^n
\end{align*}$$

(P'')

with the same optimal value as (P), and with $|J| \leq 2^n - 1$. Of course, $J$ may depend on $M$ and the choice of unnecessary constraint to relax at each stage. Now let $M$ tend to infinity through integral values.
Then some set \( J \) must occur infinitely often, and for this set \( J \) it is clear that

\[
\begin{align*}
\max a^0_i h \\
 a^j_i h \geq \beta^j, & \quad j \in J \\
 h \in \mathbb{Z}^n
\end{align*}
\]

has the same optimal value as \( (P) \). Hence Theorem 2 follows from Theorem 1.

Finally we indicate how the requirement (a) of integral data can be removed. The analyses of Sections 2 and 3 are carried out as before except that all constraints of the form \( c^0_i h \geq \epsilon^0 + l \) or \( a^i_i h \geq \gamma^i + l \) are replaced by \( c^0_i h > \epsilon^0 \) or \( a^i_i h > \gamma^i \). The programs now have mixed weak and strong inequalities, but with appropriate minor changes the arguments remain valid.

5. Comparison with Scarf's Proof

Let \( A \) be the matrix with rows \( a^0, a^1, \ldots, a^m \). Then the set of "production vectors" \( Ah^0, \ldots, Ah^q \), together with slack vectors for rows \( q+1 \) through \( m \), has much in common with a final primitive set as found in Scarf's proof [1] of Theorem 2. Specifically, if \( I = \{1, \ldots, q\} \) is a set of binding constraints for \( (P_q) \), there is no vector \( h \in \mathbb{Z}^n \) with each coordinate of \( Ah \) strictly greater than the minimum of the corresponding coordinates of the vectors above. For then we would have

\[ Ah > (\gamma^0, \ldots, \gamma^q, -M, \ldots, -M) \]

for some large \( M \) and hence

\[ Ah \geq (\gamma^0 + l, \ldots, \gamma^q + l, -M + l, \ldots, -M + l); \]

thus \( h \) would contradict the optimality of \( h^q \) in \( (P_q) \). Of course this set may well not be a primi-
tive set, for the latter requires tie-breaking rules if some coordinate of \( A \) is equal to the corresponding \( y_j \) or \(-M\). Our proof shows that such details of tie-breaking are unnecessary to prove the main result. However, the construction above of problems \((P_0), \ldots, (P_q)\) was motivated primarily by Scarf's proof and the resulting primitive set.

There is also a strong similarity in the use of the sequence of integers \( M \). The reason for its use appears to be quite different in the two approaches, however. Scarf uses the big \( M \) to bound the set of "production vectors" and enable his complementary pivoting algorithm to start. We need the big \( M \) solely to enable the results established with the boundedness condition to be used when constraints are successively relaxed.

Reference