Convex Hull Representations of the *at_least* Predicate of Constraint Satisfaction[1]

Hong Yan

The Hong Kong Polytechnic University, Hong Kong

H. P. Williams

Operational Research Group,
Department of Management,
London School of Economics and Political Science

Working Paper LSEOR 07.97 ISBN: 978-0-85328-053-8

[1] The research is partially supported by The Hong Kong UGC CERG Research Grant B-Q00U and The Hong Kong Polytechnic University Research Grant A-PA97.

First published in Great Britain in 2007
by the Operational Research Group, Department of Management
London School of Economics and Political Science

Typeset, printed and bound by:

The London School of Economics and Political Science
Houghton Street
London WC2A 2AE

Working Paper No: LSEOR 07.97 ISBN: 978-0-85328-053-81

Convex Hull Representations of the *at_least* Predicate of Constraint Satisfaction[2]

Hong Yan
The Hong Kong Polytechnic University, Hong Kong
H. P. Williams
London School of Economics, UK

May, 2007

Abstract

The predicate "$at\_least_m\{x_1,\cdots,x_n\}=k$" says that ``at least *m* of x_i 's take value *k*. This, and its generalisations, is one of the most commonly used predicates in Constraint Satisfaction. The structure of its feasibility set is exhibited. The facet defining constraints for the convex hull are then described and proved. The differences of the convex hulls are shown when the variables are or are not bounded above.

Keywords: Constraint Satisfaction, Constraint Logic Programming, *at_least* Predicate, Integer programming, Facets, Convex Hull, MIP representability.

1. Introduction

In [4], we described a tight linear inequality representation of the logic predicate $all\_different$, which says that all variables from a given set always take different values. In this work, we discuss another important predicate. Given a set of variables $\{x_1,x_2\cdots,x_n\}$, where x_i 's take non-negative integer values, we consider a predicate "at least *m* of x_i 's take a particular value *k*". This can be denoted as

$$at\_least_m\{x_1,x_2\cdots,x_n\}=k \tag{1}$$

Denote the set of all points satisfying (1) by S. We point out that unless finite upper bounds are placed on the x_i 's the condition is not representable by linear inequalities. We denote the convex hull of S by $conv(S)$. We identify all facets of $conv(S)$.

[2] The research is partially supported by The Hong Kong UGC CERG Research Grant B-Q00U and The Hong Kong Polytechnic University Research Grant A-PA97.

This is a generalization of the important logical condition expressed when $k=1$ and the x_i are 0-1 variables expressing a lower bound on the cardinality of a set. It was shown by McKinnon and Williams [2] that this predicate plays a fundamental role in integer programming formulation, if one allows nesting of the at_least condition (expressed in that paper as "$greater\_than\_or\_equal\_to$"), i.e. the truth of the predicate is itself equivalent to another 0-1 variable taking the value 1.

By exploring such "building blocks" of integer programming formulation and identifying their convex hull, we aim to provide a framework for obtaining tighter linear programming relaxations for practical integer programming formulations.

The paper is arranged as follows: Section 2 gives a simple example of the at_least predicate. It discusses two important cases: when the variables are not bounded above and when the variables are bounded above. For these cases, the convex hull descriptions have some significant difference. Section 3 focuses on the convex hull representation of the predicate "at_least" with all unbounded variables. Section 4 concentrates on the convex hull description of the predicate with all variables bounded above. It is clear it would be easy to obtain the similar results when only a partial set of variables are bounded and others are not. Section 5 summarizes the research.

2. Illustrative Examples

We discuss the convex hull of an "at_least"' predicate in two cases: when the variables are not bounded above, and when the variables are bounded above. In the first case, all variables will take any non-negative integer value. It is then clear that S and $conv(S)$ are unbounded. In the second case, without loss of the generality, we assume that all variables are bounded by the same value. The following two simple examples provide a preliminary understanding on the structure of the "at_least" and on the $conv(S)$.

Consider predicate $at\_least_1\{x_1, x_2\} = 2$. In the first case, let both x_1 and x_2 be bounded by a positive integer L, $x_1 \le L$ and $x_2 \le L$. Then it is clear that $conv(S)$ is given by

\begin{eqnarray}

$$
\begin{aligned}
x_1 + x_2 &\geq 2 \\
x_1 + x_2 &\leq L + 2 \\
x_1 - (\frac{L}{2} - 1)x_2 &\leq 2 \\
-(\frac{L}{2} - 1)x_1 + x_2 &\leq 2
\end{aligned} \tag{2}
$$

and is illustrated by Figure 1, where all the poins in S are located on the dotted lines.

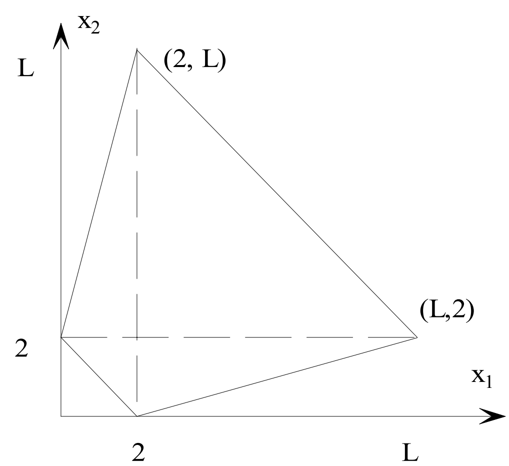

Figure 1. $conv(S)$ **when** x_1 **and** x_2 **are bounded.**

In the second case, both x_1 and x_2 are unbounded above. Then it is easy to see that $conv(S)$ is given by

$$
\begin{aligned}
x_1 + x_2 &\geq 2 \\
x_1 &\geq 0 \\
x_2 &\geq 0
\end{aligned} \tag{3}
$$

and is illustrated by Figure 2.

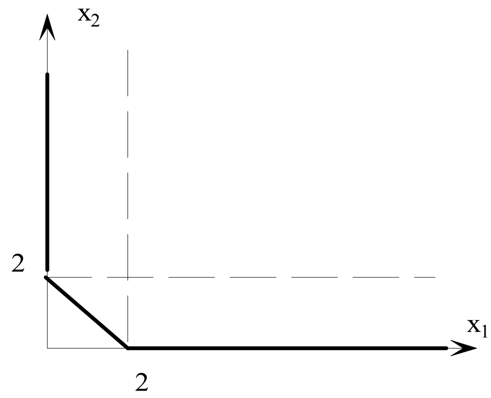

Figure 2. $conv(S)$ **when** x_1 **and** x_2 **are unbounded.**

In Figure 2, the feasible set is given by the two dotted lines. Each of these two lines is a polytope. But each polytope has a different recession direction. It is proved by Jeroslow [1] that in this case the union of the polytopes is "non mixed integer programming representable". That is, it cannot be represented as the feasible set of a finite collection of inequalities in continuous and integer variables. Therefore, we cannot represent the set as the feasible region of a (mixed) integer programme. In practice of course one can always find finite upper bounds that apply to the variable x_i, but it is important to note the theoretical impossibility of representing the condition otherwise. However, it is still possible to define its convex hull as the minimum convex set containing the feasible solutions. Such a convex hull is unbounded. Perhaps it is interesting to see that (2) can be reduced to (3) by letting $L \to \infty$.

3. Convex Hull Representation of Predicate "at_least" with Unbounded Variables

The following lemma is extended from Lemma 1 in [4].

Lemma 1 Given a vector

$$(a_1, a_2 \cdots, a_n),$$

where not all components a_i 's are the same and they are arranged in increasing order $a_1 \le a_2 \le \cdots \le a_n$, and a vector

$$(b_1, b_2, \cdots, b_n),$$

where not all components b_i 's are the same. Rearrange components b_i 's in decreasing order,

$$b_1^{'} \geq b_2^{'} \geq \cdots \geq b_n^{'}, \tag{4}$$

Then

$$\sum_{i=1}^{n} a_i b_i^{'} < \sum_{i=1}^{n} a_i b_i. \tag{5}$$

Proof: The result is shown by induction. For $n = 2$, since we assumed that not all a_i 's are the same and not all b_i 's are the same, then we have

$$a_1 < a_2, \quad \text{and} \quad b_2 > b_1.$$

Thus

$$(a_1 b_2 + a_2 b_1) - (a_1 b_1 + a_2 b_2) = (a_1 - a_2)(b_2 - b_1) < 0$$

Assume that (5) holds for $n-1$. Rearrange b_i 's such that (4) holds. If $b_n^{'} = b_n$, then (5) holds for n. Assume that

$$b_t^{'} = b_t, \qquad \forall t \leq n-1.$$

Then

$$\begin{aligned} \sum_{i=1}^{n} a_i b_i^{'} &= \sum_{i=1}^{n-1} a_i b_i^{'} + a_n b_n^{'} \\ &= \sum_{i=1}^{n-1} a_i b_i^{'} + a_n b_t \\ &< \sum_{i=1, i \neq t}^{n-1} a_i b_i + a_t b_n + a_n b_t \end{aligned}$$

Since \$ $b_n \neq b_t$, thus $b_n > b_t$. Then

$$(a_t b_n + a_n b_t) - (a_t b_t + a_n b_n) = (a_t - a_n)(b_n - b_t) \leq 0$$

Therefore,

$$\sum_{i=1}^{n} a_i b_i' < \sum_{i=1, i\neq t}^{n-1} a_i b_i + a_t b_n + a_n b_t$$
$$\leq \sum_{i=1, i\neq t}^{n-1} a_i b_i + a_t b_t + a_n b_n$$
$$= \sum_{i=1}^{n} a_i b_i .$$

Thus, (5) holds for all n. □

Lemma 2 Inequality

$$\sum_{i=1}^{n} x_i \geq mk \tag{6}$$

is a facet of $conv(S)$, the convex hull of feasible point set of (1).

Proof: First of all, it is obvious that (6) is valid for (1). Then, we show that it defines a facet by identifying n affinely independent points of (1) on (6).

We choose the first $m+1$ points as the following. Let the first $m+1$ components of x, $(x_1, \cdots, x_m, x_{m+1})$ take values

$$\begin{array}{l} (0, k, \cdots, k, k), \\ (k, 0, \cdots, k, k), \\ \cdots \\ (k, k, \cdots, 0, k), \\ (k, k, \cdots, k, 0), \end{array}$$

and other components take values 0.

The rest of the points all have the first $m-1$ components taking value k. And let other components be 0 except $x_j = k$, for $j = m+2, \cdots, n$ alternatively. That is, the n points thus obtained are listed as the following,

$$
\begin{aligned}
&(0,k,\cdots,k,k,0,0,\cdots,0),\\
&(k,0,\cdots,k,k,0,0,\cdots,0),\\
&\cdots\\
&(k,k,\cdots,0,k,0,0,\cdots,0),\\
&(k,k,\cdots,k,0,0,0,\cdots,0),\\
&(k,k,\cdots,0,0,k,0,\cdots,0),\\
&(k,k,\cdots,0,0,0,k,\cdots,0),\\
&\cdots\\
&(k,k,\cdots,0,0,0,0,\cdots,k).
\end{aligned}
$$

It is easy to see that they are affinely independent points. They are all on (6). Thus, (6) defines a facet. □

Lemma 3 For $a_i \neq 0$, $i = 1,\cdots,n,$ and $b > 0$, if

$$a_1 x_1 + \cdots + a_n x_n \geq b \tag{7}$$

defines a facet of $conv(S)$, then it is a positive scalar multiple of (6).

Proof: Let (7) define a facet of $conv(S)$. If a_i 's are different, then without loss of the generality, assume that

$$a_1 \leq a_2 \leq \cdots \leq a_n.$$

Consider a point $p^\circ = (x_1^\circ,\cdots,x_n^\circ)$ on the facet,

$$a_1 x_1^\circ + \cdots + a_n x_n^\circ = b$$

Rearrange the components of $p^\circ = (x_1^\circ,\cdots,x_n^\circ)$ in decreasing order and we have a new point p'

$$p' = (x_1',\cdots,x_n')$$

such that

$$x_1' \geq \cdots \geq x_n'.$$

Since p' is also a feasible point,

$$a_1 x_1' + \cdots + a_n x_n' \geq b \tag{8}$$

On the other hand, if not all a_i 's are the same, by Lemma 1, we have

$$a_1 x_1^{'} + \cdots + a_n x_n^{'} < a_1 x_1^{\circ} + \cdots + a_n x_n^{\circ} = b \tag{9}$$

This contradicts to (8). Thus we must have $a_i = \alpha$, for \$i $i = 1, \cdots, n$ and α is a constant. Thus (7) is rewritten as

$$\alpha(x_1 + \cdots + x_n) \geq b \tag{10}$$

It is clear that $\alpha > 0$, since otherwise (10) cannot be valid. Therefore, (10) can be written as

$$x_1 + \cdots + x_n \geq \overline{b} \tag{11}$$

where $\overline{b} = b / \alpha \geq 0$. Since (11) is a facet, $\overline{b}$ takes the smallest value, which is mk . This completes the proof. □

Note that S is an open set and unbounded above. Denote $S^{\circ} \subset S$, the points which have exactly m elements taking value k and other elements taking value 0. Let

$$I = \{1, \cdots, n\} \quad \text{and} \quad I_1 = \{i_1, \cdots, i_m\}.$$

Then, S° is denoted by

$$S^{\circ} = \{(x_1, \cdots, x_n) \mid x_i = k, \text{ for } i \in I_1, x_i = 0, \text{ for } i \in I \setminus I_1\}. \tag{12}$$

It is clear that $\forall x \in S$, there exists a point $x^{\circ} \in S^{\circ}$ such that $x \geq x^{\circ}$. That is, each element of x is no smaller than the corresponding element of x° . The facet (6) has all points in S° supported by it. Furthermore, we have the following,

Lemma 4 For $c \in R^n$, consider

$$\min\{cx : x \in S\}. \tag{13}$$

S° is defined as (12). The optimal solution of (13) is achieved on some point in S° . That is,

$$\min\{cx : x \in S\} = \min\{cx : x \in S^{\circ}\}. \tag{14}$$

Proof: Since S is not bounded above, then for any element of c , $c_i < 0$, (13) will not have an optimal solution. On the other hand, if $c \geq 0$, then from the definition of S° , it is easy to see that (14) holds. □

In this sense, we call (6) the "main facet"'. Since the tight representation of a predicate or a logical condition usually contains a large number of linear inequalities, it is always important to balance the tightness and the size of a ILP model. Therefore, sometimes we may only need to bring into the ILP model the main facet (for more details of the discussion, see [5] and [6].

Before presenting other facets of $conv(S)$, let us consider a simple predicate as follows,

$$at\_least_2\{x_1, x_2, x_3\} = 2,$$

and x_1, x_2, x_3 are unbounded. The feasible set of (15) is shown in Figure 3, with vertices represented by circles on the graph.

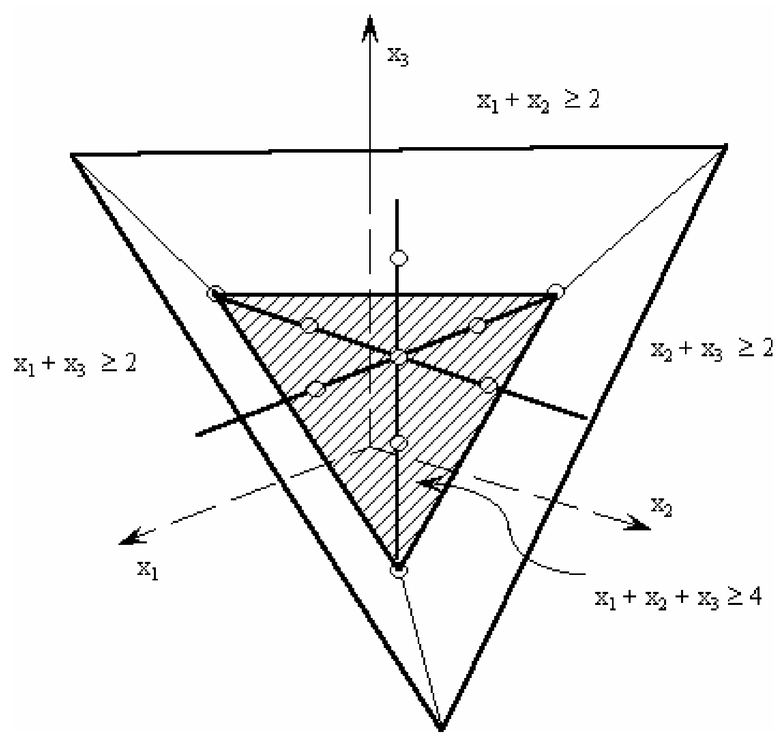

Figure 3. $conv(S)$ **for predicate (15)**

From Figure 3, it is clear that in addition to facet

$$x_1 + x_2 + x_3 \geq 4,$$

the following 3 inequalities

$$x_1 + x_2 \geq 2,$$
$$x_2 + x_3 \geq 2,$$
$$x_1 + x_3 \geq 2,$$

are all facets. In fact, for example, it is clear that $x_1 + x_2 \geq 2,$ is clearly a proper face. To show that it is a facet, we only need to identify 3 affinely independent points on it. By definition 1.4

and 2.5 of [3], it is easy to see that (2, 0, 2), (0, 2, 2) and (0, 2, 3) are affinely independent on $x_1 + x_2 = 2$. The convex hull is described by

$$\begin{aligned}
&x_1 + x_2 + x_3 \ge 4, \\
&x_1 + x_2 \ge 2, \\
&x_2 + x_3 \ge 2, \\
&x_1 + x_3 \ge 2 \\
&x_1 \ge 0, x_2 \ge 0, x_3 \ge 0.
\end{aligned}$$

In the following, we give the general description of $conv(S)$. First, we have the following

Lemma 5 For an integer t such that $n-(m-1) \le t < n$,

$$x_{i_1} + \cdots + x_{i_t} \ge (m-n+t)k, \quad \{i_1, \cdots, i_t\} \subseteq \{1, \cdots, n\} \tag{16}$$

defines a facet of $conv(S)$.

Proof: For $n-(m-1) \le t < n$, there are $n-t$ variables x_i not included in (16). Therefore, at most all these missing variables take value k. Then there are at least $m-(n-t)$ variables in (16) still taking value k. Thus, we must have

$$x_{i_1} + \cdots + x_{i_t} \ge (m-n+t)k.$$

That is, (16) is valid. To show that it is a facet, we need to identify n affinely independent points on it. Note that as shown in the previous example, these points are necessarily in S. It is clear that we can take similar points as in the proof of Lemma 2, except that we can allow those variables not included in (16) to take values different than k, for example $k+1$ or $k+2$ etc. To ensure these points are supported by (16), it is only necessary to keep $m-(n-t)$ variables involved in (16) taking value k. □

Let

$$P = \left\{ (x_1, \cdots, x_n) \middle| \begin{array}{l} x_{i_1} + \cdots + x_{i_t} \ge (m-n+t)k, \quad \{i_1, \cdots, i_t\} \subseteq \{1, \cdots, n\} \\ t = n, n-1, \cdots, n-m+1 \end{array} \right\},$$

This description of P has $2^n - 1$ constraints.

Combining lemmas above, we have,

Theorem 1 The facets of $conv(S)$ are defined by the inequalities in P. In other words,

$$P = conv(S) \tag{17}$$

□

4. Convex Hull Representation of Predicate "at_least" with Bounded Variables

This section consider the convex hull representation of the "at_least"predicate where all variables are bounded above.

It is common in constraint programming for variables to have constantly changing bounds, both lower and upper. However, the convex hull becomes very complex when the bounds are different. In practice all the upper bounds can be set equal to the largest one when generating the convex hull. Therefore, in this research, we assume that all variables have the same upper bound. That is,

$$at\_least_m\{x_1, x_2, \cdots, x_n\} = k,\ x_i \le L,\ i = 1, \cdots, n. \tag{18}$$

Again, let S denote all the points satisfying (18) and $conv(S)$ denote the convex hull of S. When all variables are bounded above, the convex hull of "at_least" is more complicated. This can be seen from the following example. Consider

$$at\_least_2\{x_1, x_2, x_3\} = 2,\ x_i \le 5,\ i = 1, 2, 3. \tag{19}$$

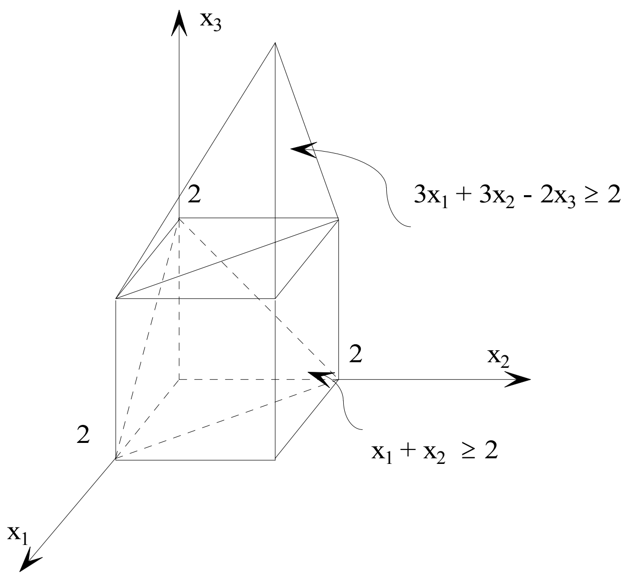

Figure 4. $conv(S)$ **for predicate (19)**

It is not difficult to see that the convex hull of predicate (19) consists of the following three parts:

$$x_1 + x_2 + x_3 \geq 4 \quad (km)$$

$$\begin{aligned} 3x_1 + 3x_2 - 2x_3 &\geq 2, \\ 3x_1 - 2x_2 + 3x_3 &\geq 2, \\ -2x_1 + 3x_2 + 4x_3 &\geq 2, \end{aligned}$$

$$x_1 + x_2 + x_3 \leq 9 \ (km + L)$$

In fact, with the similar process as in Lemma 1 in the last section, we can show the following.

Lemma 6 Inequality

$$\sum_{i=1}^{n} x_i \geq mk \tag{20}$$

defines a facet of $conv(S)$.

Furthermore, another facet can be shown by the following lemma.

Lemma 7 Inequality

$$\sum_{i=1}^{n} x_i \leq mk + (n-m)L \tag{21}$$

defines a facet of $conv(S)$.

The proof of Lemma 7 is rather lengthy, but not difficult. First, it is easy to see that (21) is valid, since at least m of $\{x_1, x_2, \cdots, x_n\}$ take value k only $n-m$ of x_i's can achieve their upper bound L. Then we can identify n affinely independent points in S such that (20) holds as equation.

The next lemma gives other facets of $conv(S)$.

Lemma 8 For $I = \{1, 2, \cdots, n\}$, denote

$$I_1 = \{i_1, i_2, \cdots, i_m\} \subseteq I; \quad I_2 = I \setminus I_1.$$

Then, $\forall I_1 \subseteq I$, inequality

$$(L-k)\sum_{i \in I_1} x_i - k\sum_{i \in I_2} x_i \geq (m-1)Lk - mk^2 \tag{21}$$

defines a facet of $conv(S)$.

Proof: First, we show that (21) is valid. Since $x_i \leq L$, $\forall i \in I$, and that at least m of x_i's take value k, the smallest that the left-hand-side can be is, by letting $x_i = L$ for $i \in I_2$ and $x_i = k$ for $i \in I_1$,

$$\begin{aligned}
(L-k)\sum_{i \in I_1} x_i - k\sum_{i \in I_2} x_i &\geq (L-k)(mk) - k(n-m)L \\
&= mLk - mk^2 - k(n-m)L \\
&= (m-1)Lk - mk^2 - (n-m-1)Lk \\
&\geq (m-1)Lk - mk^2
\end{aligned}$$

since $(n-m-1) \geq 0$ by definition.

Then, we show that (21) defines a facet by identifying n affinely independent points of S on it.

We choose the first $(n-m)$ points as the following. For all $i \in I_1$, let $x_i = k$; and for each one $i \in I_2$, $x_i = L$, and for other $i \in I_2$, let $x_i = 0$. That is,

$$
\begin{aligned}
&(k,k,\cdots,k,L,0,\cdots,0),\\
&(k,k,\cdots,k,0,L,\cdots,0),\\
&\cdots\\
&(k,k,\cdots,k,0,0,\cdots,L).
\end{aligned}
$$

It is clear that

$$(L-k)\sum_{i\in I_1} x_i - k\sum_{i\in I_2} x_i = (L-k)mk - kL = (m-1)kL - mk^2$$

The next m points are chosen as the following. For each one $i \in I_1$ let $x_i = 0$ and for other $i \in I_1$, let $x_i = k$; in all cases, for the first $i \in I_2$, let $x_i = k$ and all other $x_i = 0$. That is,

$$
\begin{aligned}
&(0,k,\cdots,k,k,0,\cdots,0),\\
&(k,0,\cdots,k,k,0,\cdots,0),\\
&\cdots\\
&(k,k,\cdots,0,k,0,\cdots,0).
\end{aligned}
$$

Again, it is clearly that

$$(L-k)\sum_{i\in I_1} x_i - k\sum_{i\in I_2} x_i = (L-k)(m-1)k - k^2 = (m-1)kL - mk^2$$

Then, the n points thus obtained are listed as rows of the following matrix A,

$$
A = \begin{Bmatrix}
k,k,\cdots,k,L,0,\cdots,0\\
k,k,\cdots,k,0,L,\cdots,0\\
\cdots\\
k,k,\cdots,k,0,0,\cdots,L\\
0,k,\cdots,k,k,0,\cdots,0\\
k,0,\cdots,k,k,0,\cdots,0\\
\cdots\\
k,k,\cdots,0,k,0,\cdots,0
\end{Bmatrix}
$$

It is not difficult to show that $\det(A) \neq 0$. Thus the n points identified above are affinely independent. Thus, (21) defines a facet. □

5. Conclusion

By understanding the facial structure of such commonly used predicates in constraint satisfaction, we hope to contribute to the efficient solving of such models by integer programming but making use of the undoubted representational strength of constraint satisfaction. Our ambition is to extend our analysis of the at_least predicate to include nesting as done by McKinnon and Williams. It would then become possible to have a recursive procedure for defining the convex hull of an integer programming problem.

Reference

[1] Jeroslow, R. (1987), "Representability of mixed integer programming", Discrete Applied Mathematics, 17, 223-243.

[2] McKinnon, K.I. and H.P. Williams (1989), "Constructing integer programming models by the predicate calculus", Annals of Operations Research, 21, 227-246.

[3] Nemhauser, G.L. and L. A. Wolsey (1988), Integer and Combinatorial Optimization, John Wiley & Sons, Inc. New York.

[4] Williams, H.P. and H. Yan (1999), "Representations of the *all_different* predicate constraint satisfaction in integer programming", INFORMS Journal of Computing, 13(2), 96-103.

[5] Yan, H. (1998). "An ILP model with logical constraints for mail sorting facility selection", Journal of Operational Research Society, 49, 273-277.

[6] Yan, H. and J.N. Hooker (1999), "Tight representation of logic constraints as cardinality rules", Mathematical Programming, 85(2), 363-377.