ALTERNATIVE GROUP RELAXATION OF INTEGER PROGRAMMING PROBLEMS

P. BERTOLAZZI, C. LEPORELLI and M. LUCERTINI
Istituto di Automatica and IASI of C.N.R., Via Eudossiana 18, 00184 Rome, Italy

Received 30 January 1981
Revised 29 July 1982 and 5 July 1983

The classical group approach to integer linear programming problems (IP) can be generalized in order to obtain group minimization problems with different computational load and different relaxation.

The aim of this work is to analyze some group problems, associated to the same (IP), both from the point of view of the relaxation of the (IP) and of the complexity of the group solution algorithm: evaluation criteria for these group problems are pointed out.

1. Introduction

The group theoretical approach for solving Integer Programming (IP) problems introduces a relaxation of the original IP problem; more precisely IP is transformed in an integer programming problem over a cone (IPC) [2,4,10]. The tightness of the relaxation introduced and the size of the group are, generally speaking, conflicting goals (i.e. as the size gets smaller the relaxation gets weaker and the group problem is less likely to represent the IP problem).

In the classical group approach the feasible region obtained from IP, dropping the integrality constraints, is usually replaced by the smallest cone containing the region and with the vertex in the LP optimal solution.

It should be noticed that the solution of the group problem (GP) may not only be useful in finding the optimal solution of IP but can also be useful as a lower bound in a branch and bound procedure. In fact if the IPC solution is not feasible for IP, the corresponding objective function is a lower bound of the optimal solution, tighter than the LP one.

The size of the group corresponding to this relaxation is less or equal to the determinant of the optimal basis matrix of the LP problem obtained from IP dropping the integrality constraints.

The computational time needed to solve GP generally increases with the size of the group [8,9,10]. In fact GP is an NP-complete problem but can be solved in pseudopolynomial time $O(n \cdot a_{\text{max}}^3)$, where $n$ is the number of variables and $a_{\text{max}}$ is the largest entry (in absolute value) that origine IP [12].

One approach to overcome the computational problems is proposed in [9,10],
where more efficient algorithms to solve GP without modifying the size of the group are presented.

Another approach consists in reducing the size of the group, allowing a weaker relaxation.

In order to reduce the size of the group many procedures have been proposed [10], all of them based essentially on the reduction of the determinant of the basis matrix. For example [6] presents number-theoretic procedures based on dividing for a suitable constant all the coefficients of a constraint; this procedure reduces the value of the determinant and thus generally the size of the group.

On the other hand, Jeroslow [7] has shown some generalizations of the usual group problem construction, that suggest a different way of reducing the size of the group. Given an IP problem, let \( \{A\} \) indicate the group generated by the columns of the matrix \( A \) of equality constraints of IP. The corresponding GPH problem is the minimum cost problem on the finite quotient group \( (\{A\} : \{H\}) \), where \( H \) is an \((m \times t)\) matrix of full rank \( m (m \leq t) \) whose columns are elements of \( \{A\} \) (and \( \{H\} \) is the subgroup of \( \{A\} \) generated by \( H \)). Of course the size of the group depends on how the matrix \( H \) spans \( A \). The classical group approach gives \( H = B \), where \( B \) is the optimal LP basis.

The GPH problem can be written as a set of congruence constraints. Generally these congruence constraints are expressed utilizing the Smith Normal Form of \( H \). As stated earlier the GPH problem is in general a relaxation of the IP. For example if \( H \) is a submatrix of \( A \) the relaxation corresponds to dropping the nonnegativity constraints on the variables associated to the columns of \( H \). A simple way of defining a group \( \{H\} \) that strictly contains \( \{B\} \), is to take \( H \) as \([B : h]\) when \( h \) is a nonbasic column of \( A \) such that \( h \notin \{B\} \).

The properties of the problem deriving from this definition of \( H \), dropping the nonnegativity constraints on \( x_h \), are analyzed in Section 2. In Section 3 an approach based on reintroducing the dropping constraint by a Lagrangian method is presented.

The main result of the paper consists in finding a way to reduce the size of the group with an 'a priori' bound of the distance between the obtained value of the objective function and the optimal solution of the original GP. It is interesting to note that under some conditions the bound is zero, i.e. the two solutions coincide.

2. A reduced group problem

The problem IP can be written in the following form

\[
\min (z - z^*_{LP}) = c_h x_h + c_N x_N,
\]

\[(IP) \quad \Delta ' x_B + h x_h + N x_N = b,
\]

\[C' x_B, x_h, x_N \geq 0, \text{ integers.}\]

where \( \Delta ' \) is the (diagonal) optimal LP basic matrix expressed in Smith normal form
Alternative group relaxation

\((\Delta' = R'BC')\), \([h : N]\) is the corresponding nonbasic matrix, \(c_h \geq 0, c_N \geq 0\) are the optimal LP costs \(z^*_L\) is the optimal value of the LP associated to IP, \(R'\) and \(C'\) are integer unimodular \((m \times m)\) matrices and \(C'x_B\) are the original variables of the problem.

The classical group problem GP, formulated with respect to \(B\), expressed in Smith Normal Form [10] is

\[
\min z_G = c_h x_h + c_N x_N,
\]

\((\text{GP})\)

\[
[h : N] \begin{bmatrix} x_h \\ x_N \end{bmatrix} \equiv b \pmod{\delta'}
\]

\(x_h, x_N \geq 0\), integers

where \(\equiv\) indicates a congruence relation, \(\delta'_i = \Delta'_i\) and \(\prod_i \delta'_i = |\det B| = \det \Delta'.\) Let \((z^*_G, x^*_h, x^*_N)\) be the optimal solution of GP. The reduced group problem GPH, formulated with respect to \(H = [\Delta' : h]\), expressed in Smith Normal Form is

\[
\min z_H = \tilde{d} + \tilde{c}_h x'_h + \tilde{c}_N x_N,
\]

\((\text{GPH})\)

\(RNx_N \equiv Rb \pmod{\delta},\)

\(x_N \geq 0\), integer, \(x'_h\) integer

where \(R\) and \(C\) are integer unimodular matrices and (see [11]):

\[
\begin{bmatrix}
\delta_1 & 0 & \cdots & 0 \\
& \ddots & \vdots & \vdots \\
0 & \cdots & \cdots & \delta_m \\
\end{bmatrix} = [\Delta : 0] = R[\Delta' : h]C,
\]

\(\det \Delta = \prod_i \delta_i = \gcd\{\text{determinants of order } m \text{ minors of } H\}
\]

\(= \gcd\{\det \Delta', (\det \Delta')h_1/\delta'_1, \ldots, (\det \Delta')h_m/\delta'_m\}.
\)

Remark that the computation of the new determinant involves only \(m\) products and divisions and the computation of the gcd of \(m\) integers and does not require the computation of the new Smith Normal Form.

\[
\begin{bmatrix}
x_B \\
x_h \\
\end{bmatrix} = \begin{bmatrix}
C_{11} & C_{12} \\
C_{21} & C_{22} \\
\end{bmatrix} \begin{bmatrix}
x_B' \\
x_h' \\
\end{bmatrix},
\]

\(d = c_h C_{21} \Delta^{-1} Rb, \quad \tilde{c}_h = c_h C_{22}, \quad \tilde{c}_N = c_N - c_h C_{21} \Delta^{-1} R N.
\)

Lemma 1. If \(\det C = 1\), then \(C_{22} = (\det \Delta')/(\det \Delta)\) and \(1 \leq C_{22} \leq \delta'_m\).

Proof. This result follows directly from (2.1) and from the expression of the determinant of a partitioned matrix. \(\square\)

The optimal solution of GPH is in general unbounded. In fact, if \(\tilde{c}_h \neq 0\), then, as \(x'_h\) is only constrained to be integer, \(z^*_H = -\infty\); on the other hand if some entries of \(\tilde{c}_N\) are negative also \(z^*_H = -\infty\).
However, supposing that $x_N \geq 0$, it is possible to give an approximate algorithm to solve GP as follows (let $\lceil x \rceil$ indicate the smallest integer greater or equal to $x$):

**Algorithm 1.**

**Step 1.** Solve GPH with $x_N = 0$; let $x_N$ be the solution.

**Step 2.** Given $x_N$, find $x'_N$ such that the constraint

$$x'_N = C_{21}x'_B + C_{22}x' = C_{21}A^{-1}R(b - Nx_N) + C_{22}x'_N \geq 0$$

(2.2)

is satisfied; the value of $x'_N$, that minimize $z_H$, given (2.2), is

$$x'_N = \lceil \Gamma(x_N) \rceil, \quad \Gamma(x_N) = \frac{C_{21}A^{-1}R(Nx_N - b)}{C_{22}}$$

(2.3)

and then

$$x_N = C_{22}(\lceil \Gamma(x_N) \rceil, \quad \Gamma(x_N)).$$

This algorithm gives a feasible solution of GP ($x'_N, x_N$) with objective function

$$z_G = c_h x'_h + c_N x_N = d + c_h x'_h + c_N x_N.$$

The additional computations required by Algorithm 1 are mainly due to the computation of the new Smith Normal Form. Remark that the following nonlinear group problem is equivalent to GP:

$$\min z_N = d + c_h \lceil \Gamma(x_N) \rceil + c_N x_N, \quad RNx_N \equiv Rb \pmod{\delta}, \quad x_N \geq 0, \text{ integer}.$$  

3. A Lagrangian approach to solve GP

3.1. The constraint $x_N \geq 0$ (see (2.2)) can be reintroduced in the objective function of GPH using Lagrangian technique [14]; in this way a tighter relaxation of GP is obtained.

The new problem (GPL) can be written:

$$(GPL) \quad L(\lambda) = \min z_L = z_H + \lambda [C_{22}x'_N - C_{21}A^{-1}R(Nx_N - b)], \quad RNx_N \equiv Rb \pmod{\delta}, \quad x_N \geq 0, \text{ integer}$$

where $\lambda \geq 0$ is a scalar, and $x'_N$ doesn't appear in the constraints of GPL, but is only constrained to be integer.

Obviously the optimal value of the Lagrangian is given by

$$L^* = \max_{\lambda \geq 0} L(\lambda).$$
Remark that
\[ z_{L}^{*} \leq L^{*} \leq z_{G}^{*} \]
where \( z_{L}^{*} \) indicates the optimal solution of the linear programming problem associated to IP.

**Lemma 2.** A necessary optimality condition for GPL is \( \lambda = c_{h} \).

**Proof.** In fact \( c_{h}C_{22} - \lambda C_{22} \) is the cost coefficient of \( x_{h}^{'} \) in GPL. As \( x_{h}^{'} \) doesn’t appear in the constraints if the cost coefficient is different from zero the objective function is unbounded. From \( C_{22} \leq 1 \) the result follows. \( \square \)

By substituting the optimal value of \( \lambda \) in \( z_{L} \), we obtain
\[
L(\lambda) = \min z_{L} = c_{N}x_{N}
\]
where \( x_{N} \) belongs to the feasible region of GPH, and \( x_{h}^{'} \) is not constrained and doesn’t affect the objective function of GPL. Let \( (L^{*}, x_{N}^{L}) \) be the optimal solution of GPL.

**Lemma 3.** \( (x_{h}^{*}, x_{N}^{L}) \) leads to a feasible solution \( (x_{h}, x_{N}^{L}) \) of GP if (see (2.2), (2.3))
\[
x_{h}^{*} \geq \Gamma(x_{N}^{L}). \quad \text{integer.} \quad \square
\]

Obviously, as \( c_{h}C_{22} \geq 0 \), among these feasible solutions, the best one is obtained from:
\[
\hat{x}_{h} = \left[ \Gamma(x_{N}^{L}) \right] \quad (\hat{x}_{h} = C_{22}(\left[ \Gamma(x_{N}^{L}) \right] - \Gamma(x_{N}^{L}))).
\]

It is possible now to give an approximate algorithm as follows:

**Algorithm 2.**

*Step 1.* Solve GPL, let \( x_{N}^{L} \) be the solution.

*Step 2.* Given \( x_{N}^{L} \), calculate \( \hat{x}_{h}^{'} \) (or \( \hat{x}_{h} \)); let \( \xi_{G} \) be the corresponding value of the objective function
\[
\xi_{G} = \hat{d} + \hat{c}_{h}\hat{x}_{h}^{'} + \hat{c}_{N}x_{N}^{L} = c_{h}\hat{x}_{h} + c_{N}x_{N}^{L}.
\]

3.2. A sufficient optimality condition can now be obtained with the following consideration:

(1) \( L^{*} \leq z_{G}^{*} \).

(2) From \( x_{N}^{L} \) it is possible to construct a feasible solution of GP \( (\hat{x}_{h}^{*}, x_{N}^{L}) \), see Lemma 3.

(3) The value of \( z_{G} \) associated with this feasible solution \( \xi_{G} \) is always not better than \( z_{G}^{*} \) (\( z_{G}^{*} \leq z_{G} \)).

(4) If \( L^{*} = \xi_{G} \), then the feasible solution is optimal for GP \( (x_{N}^{L} = x_{N}^{*}, \hat{x}_{h} = x_{h}^{*}) \).
Theorem 1. The maximum value of the difference \((\hat{z}_G - z^*_G)\) is given by

\[ c_h \hat{x}_h = c_h C_{22} \left( \Gamma(x^*_G) \right) - \Gamma(x^*_G) \leq c_h (C_{22} - 1) \leq c_h (d' - 1). \]

Proof. As \((\hat{z}_G - z^*_G) \leq (\hat{z}_G - L^*)\), from the expression of \(z^*_G\) and \(L^*\) follows the result; remark that the value of \((\Gamma(x^*_G) - x)\) belongs to the interval \([0, 1)\).

Corollary 1. If \(c_h \hat{x}_h = 0\), i.e. \(\Gamma(x^*_G)\) is integer, or \(c_h = 0\), then \(\hat{z}_G = z^*_G\), and \((\hat{x}_h, x^*_G)\) is the optimal solution of GP.

The result of Corollary 1 leads to a fairly generalization of the classical group approach. In fact, let \(B\) be the optimal LP basis of a given IP problem and \(H\) the set of nonbasic variables with \(c_i = 0\) \((i \in H)\) in the optimal solution. Following the classical group approach there exists (in the worst case) \(\left(\frac{|B|}{|H|}\right)\) possible equivalent basis matrices and hence the same number of group generators; all the equivalent bases lead to the same solution of the group problem (the same value of the objective function) but the group corresponding to different bases has in general a different size.

As a consequence of Corollary 1 it follows that we can obtain the same solution utilizing, as group generator, the set \(BUH\); the corresponding group size is in this case the greatest common divisor of the determinants of all the equivalent basis matrices, hence considerably smaller of all the determinants of the single basis matrices. However, it must be noticed that if some \(c_i\) are equal to zero the efficiency of the group approach is in general reduced, because of the possibility to find negative solutions for the basis variables. Obviously the reduction procedures does not affect in any way the probability of negative solutions.

3.3. It is now possible to give a procedure of choosing the column \(h\) in order to satisfy (if possible) a given level of approximation. The procedure is based on the following result.

Let \(\alpha\) be the maximum accepted value of \((\hat{z}_G - z^*_G)\) and \(C_{22}(h)\) the value of \(C_{22}\) corresponding to a given column \(h\) (remember that \(C_{22} = (\text{det } A')/\text{det } A)\):

Theorem 2. If \(C_{22}(h) \leq (\alpha + c_h)/c_h\), then \((\hat{z}_G - z^*_G) \leq \alpha\).

Proof. \((\hat{z}_G - z^*_G) \leq (\hat{z}_G - L^*) \leq c_h (C_{22}(h) - 1)\), then if \(c_h (C_{22}(h) - 1) \leq \alpha\), then \((\hat{z}_G - z^*_G) \leq \alpha\).

Algorithm 3.

Step 1. Given \(\alpha\), find \(h\) such that the equation

\[ C_{22}(h) \leq (\alpha + c_h)/c_h \]
Alternative group relaxation

is satisfied; let \( h^\star \) be the result (if no \( h^\star \) exists it means that it is not possible to guarantee 'a priori' the given approximation).

**Step 2.** Calculate \( (x_k^\star, \hat{x}_h^\star) \) utilizing Algorithm 2; \( (c_h^\star, \hat{x}_h^\star) \) is the maximum value of the error (in general less than \( [c_h^\star(C_{22}(h) - 1)] \)).

Note that the same procedure can be applied if we fix the maximum percentage error \( \alpha^\prime = ((z_G - z_{G^\star})/z_{G^\star}) \) utilizing instead of \( z_{G^\star} \) the optimal value of the LP associated to the IP (hypothized positive).

4. Numerical results

The Algorithm 2 has been tested on several problems with both cyclic and non-cyclic groups. In the latter case the bound on the group reduction given in lemma 1 guarantees that, also in very small problems, \( (\det \Delta) \) doesn't collapse to 1; consequently the problem relaxation is not too big and the algorithm performance is quite good: so, in the example 14.3 of [15], with both choices of the \( h \)-variable the algorithm gives the optimum solution of GP with a determinant reduction of respectively 2 and 6 times. In cyclic group problems (2.1) becomes

\[
\det \Delta = \delta_m = \gcd(\delta_m^\prime, h_m)
\]

and allows a quick computation of the a priori bound in Theorem 1 for several choices of the \( h \)-variable. On the other hand, in very small test problems it is easy to obtain large reduction of \( (\det \Delta) \) and so poor bounds on the difference \( z_G - z_{G^\star} \); moreover, if \( \det \Delta = 1, L^\star = z_{L^\star}^\star = 0 \).

The test problem in Table 1 has been devised to obtain some insight on the algorithm performance over a large range of group reductions. A single constraint problem with 11 nonbasic variables is considered (see the tableau on the top of the table). The standard group problem has been solved first. Then eleven experiments have been performed choosing each nonbasic variable as \( h \)-variable. The table shows for each experiment the group reduction factor, the value of the objective function and of the variables for the feasible solution of GP obtained with Algorithm 2, the associated Lagrangian lower bound on the optimal solution of GP, a priori and a posteriori bound on the difference \( z_G - z_{G^\star} \) and the actual difference in absolute and percentage terms.

In the experiments 6, 9, 10 and 11 a negative value of \( x_B \) is obtained with the proposed rule to compute \( x_h^\prime \). In these cases alternative solutions are derived improving \( x_B \geq 0 \) instead of \( x_h \geq 0 \) in the computation of \( x_h^\prime \).

The a priori bound \( c_h(C_{22}(h) - 1) \) is reasonably correlated to the actual value of \( z_G - z_{G^\star} \) and gives a good guide in the computationally inexpensive choice of \( h \). In fact choosing \( h = 1 \), in correspondence with the minimum value of the bound, the optimal solution of GP is obtained with an half-size group.

In [16] a computer program for the algorithms proposed in this paper can be
<table>
<thead>
<tr>
<th>Probl.</th>
<th>Ob. Fun.</th>
<th>Lagr.</th>
<th>$c_n x_h$</th>
<th>$c_n (C_{22} - 1)$</th>
<th>$\xi_n - \xi_h$</th>
<th>Worst. Reduc.</th>
<th>$x_B$</th>
<th>$x_N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orig.</td>
<td>17377</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>17377</td>
<td>7377</td>
<td>10000</td>
<td>10000</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>20670</td>
<td>670</td>
<td>20000</td>
<td>20000</td>
<td>3293</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>35354</td>
<td>5354</td>
<td>30000</td>
<td>40000</td>
<td>17977</td>
<td>5</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>47358</td>
<td>7358</td>
<td>40000</td>
<td>60000</td>
<td>29981</td>
<td>7</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>72798</td>
<td>2798</td>
<td>70000</td>
<td>100000</td>
<td>55421</td>
<td>11</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>110000</td>
<td>0</td>
<td>110000</td>
<td>134000</td>
<td>92623</td>
<td>105</td>
<td>-1</td>
<td>11</td>
</tr>
<tr>
<td>6*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>105</td>
<td>15</td>
<td>-94</td>
</tr>
<tr>
<td>7</td>
<td>48422</td>
<td>0</td>
<td>48422</td>
<td>71579</td>
<td>31045</td>
<td>2310</td>
<td>0</td>
<td>1562</td>
</tr>
<tr>
<td>8</td>
<td>47486</td>
<td>62</td>
<td>47486</td>
<td>69970</td>
<td>30109</td>
<td>231</td>
<td>0</td>
<td>1562</td>
</tr>
<tr>
<td>9</td>
<td>100675</td>
<td>670</td>
<td>100005</td>
<td>280014</td>
<td>83298</td>
<td>15</td>
<td>-2</td>
<td>5</td>
</tr>
<tr>
<td>9*</td>
<td></td>
<td>15</td>
<td>6</td>
<td></td>
<td></td>
<td>15</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>681374</td>
<td>1340</td>
<td>680034</td>
<td>680034</td>
<td>663997</td>
<td>35</td>
<td>-11</td>
<td>34</td>
</tr>
<tr>
<td>10*</td>
<td></td>
<td>35</td>
<td>1</td>
<td></td>
<td></td>
<td>4</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>592862</td>
<td>62</td>
<td>592800</td>
<td>598000</td>
<td>575485</td>
<td>231</td>
<td>-24</td>
<td>2</td>
</tr>
<tr>
<td>11*</td>
<td></td>
<td>231</td>
<td>1</td>
<td></td>
<td></td>
<td>2</td>
<td>2</td>
<td>-3</td>
</tr>
</tbody>
</table>
found, with an analysis of the numerical results for several test problems (Zionts [15], Haldi, IBM [17]) and problems arising from some applications.

5. Conclusions

In this paper some algorithms are proposed for the exact or approximated solution of an IP problem over a cone. In particular some sufficient conditions such that Algorithm 2 leads to an optimal solution, are given. It should be noticed that if some of nonbasic variables of the LP optimal solution have zero cost coefficient, it is always possible to find the optimal solution utilizing one of these columns as column $h$. A method of finding approximated solutions with a given level of accuracy, is also briefly presented; the crucial point of the method is a fast computation of $C_{22}(h)$.

References