Complementary approaches for the computation of the independent number of a graph

CARMO P. BRÁS
CMA, Faculdade de Ciências e Tecnologia
Universidade Nova de Lisboa
Departamento de Matemática
Campus de Caparica
2829-516 Caparica
PORTUGAL
mb@fct.unl.pt

JOAQUIM J. JÚDICE
Faculdade de Ciências e Tecnologia
Universidade de Coimbra
Departamento de Matemática and
Instituto de Telecomunicações
Largo D. Dinis, Apartado 3008
3001-454 Coimbra
PORTUGAL
joaquim.judice@co.it.pt

Abstract: The problem of finding the independent number of an undirected graph is formulated as two equivalent Mathematical Programs with Linear Complementarity Constraints ($MPLCC$). A multistarting Lemke's method is introduced for dealing with the first formulation and is able to find a good approximate of the independent number in a finite number of iterations. A sequential complementary algorithm is also discussed for the second formulation and can find the independent number at least in theory. Some computational experience is included to highlight the efficacy of the complementary approaches for computing the independent number of graphs from the Dimacs collection.

Key–Words: Linear Complementarity Problems, Graph Theory, Nonlinear Programming, Global Optimization

1 Introduction

Let $G(V, E)$ be an undirected graph, where $V = \{1, \ldots, n\}$ and E are the sets of nodes and edges, respectively. A Maximum Independent Set (MIS) of V [6] is a set $S \subseteq V$ of nodes of largest cardinality that are pairwise nonadjacent. The cardinal $|S|$ of such a set is called the Independent Number of G and is denoted by $\alpha(G)$. As discussed in [3] an MIS can be obtained by finding a global minimum of a nonconvex quadratic program with nonnegative variables [3] or by processing a nonconvex standard quadratic program [3].

Several approaches have been proposed in the literature for computing an MIS, which search such a set by solving one of the formulations discussed above [3, 6, 13]. In particular, a complementary pivoting approach based on Lemke's method [12] has been discussed in [13]. In this paper, we also investigate the use of complementary techniques for finding an MIS. To do this, we first introduce two equivalent formulations of the problem as a Mathematical Programming Problem with Linear Complementarity Constraints ($MPLCC$).

A multistarting Lemke's method and a sequential complementary (SC) algorithm [11] are then proposed to solve these two $MPLCC$ formulations. Computational experience with some graphs from the Dimacs collection is reported to highlight the efficiency and efficacy of these complementary approaches to find an MIS in practice.

The organization of the paper is as follows. The $MPLCC$ formulations are introduced in section 2. The multistarting Lemke's method and the SC algorithm are discussed in sections 3 and 4. Computational experience and some conclusions are reported in the last section.

2 Linear complementarity formulations

As discussed in [3], the problem of finding a Maximum Independent Set (MIS) of an undirected graph $G = (V, E)$ and the corresponding Independent Number $\alpha(G) = |S|$ can be stated as the following quadratic program (QP)

$$QP_1: \quad \begin{array}{ll} Maximize & e^T x - \frac{1}{2} x^T (A + I) x = \frac{\alpha(G)}{2} \\ subject\ to & x \geq 0 \end{array}$$

where A is the adjacency matrix associated to G, I is the identity matrix of order n and $e \in \mathbb{R}^n$ is a vector of ones. The Karush-Kuhn-Tucker (KKT) conditions for QP lead to the following Linear Complementarity

Problem (LCP)

$$\begin{array}{l} w = -e + (A+I)x \\ x \geq 0, w \geq 0 \\ x^T w = 0, \end{array} \tag{1}$$

where $w \in \mathbb{R}^n$ is the vector Lagrange multipliers associated to the nonnegative constraints $x \geq 0$. If x is a solution of this $LCP,$ then

$$\frac{1}{2}\,\alpha(G) = \frac{1}{2}\,e^T x - \frac{1}{2}\,x^T(-e + (A+I)x) = \frac{1}{2}\,e^T x.$$

Therefore an MIS and $\alpha(G)$ can be computed as a global minimum of the following Mathematical Program with Linear Complementarity Constraints ($MPLCC$)

$$\begin{array}{lll} MPLCC_1 : & Maximize & e^T x \\ & subject\ to & w = -e + (A+I)x \\ & & x \geq 0,\ w \geq 0 \\ & & x^T w = 0. \end{array}$$

Furthermore, it is easy to show that each solution of the LCP (1) satisfies $x_i \in \{0,1\},\ i = 1,\ldots,n$. The nodes of each independent set of G are associated with the positive components of each LCP solution. So an MIS is a set $S = \{i_1,\ldots,i_k\} \subset \{1,\ldots,n\}$ such tat the vector $\bar{x}$ defined by

$$\bar{x}_i = \begin{cases} 1 & \text{if } i \in S, \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

is a global minimum of $MPLCC_1$. Furthermore $\alpha(G) = e^T\bar{x}$.
Another quadratic programming formulation for an MIS has been proposed in [3] and is given below

$$\begin{array}{lll} QP_2 : & Minimize & \frac{1}{2}\,x^T(A+I)x \\ & subject\ to & e^T x = 1 \\ & & x \geq 0. \end{array}$$

The KKT conditions for QP_2 constitute the following Mixed LCP

$$\begin{array}{l} w = (A+I)x - ue \\ e^T x = 1 \\ x \geq 0, w \geq 0 \\ x^T w = 0, \end{array} \tag{3}$$

where $w \in \mathbb{R}^n$ and $u \in \mathbb{R}$ are the Lagrange multipliers associated to the constraints $x \geq 0$ and $e^T x = 1$ respectively. For each solution of the Mixed LCP (3)

$$\frac{1}{2}x^T(A+I)x = \frac{1}{2}x^T(w + ue) = \frac{u}{2}$$

and QP_2 is equivalent to the following $MPLCC$

$$\begin{array}{lll} MPLCC_2 : & Minimize & u \\ & subject\ to & w = (A+I)x - ue \\ & & e^T x = 1 \\ & & x \geq 0, w \geq 0, u \geq 0 \\ & & x^T w = 0. \end{array}$$

Next, we establish an interesting relationship between the constraints sets of these two $MPLCCs$.

Theorem 1 *The vector $\bar{x}$ is a solution of LCP (1) if and only if $\left(\tilde{x} = \frac{\bar{x}}{e^T\bar{x}}, \tilde{u} = \frac{1}{e^T\bar{x}}\right)$ is a solution of Mixed LCP (3).*

Proof: If $\bar{x}$ is a solution of LCP (1), then $0 \neq \bar{x} \geq 0$ and $e^T\bar{x} > 0$. Furthermore

$$\frac{\bar{w}}{e^T\bar{x}} = -\frac{e}{e^T\bar{x}} + (A+I)\left(\frac{\bar{x}}{e^T\bar{x}}\right)$$

and $\left(\tilde{w} = \frac{\bar{w}}{e^T\bar{x}}, \tilde{x} = \frac{\bar{x}}{e^T\bar{x}}, \tilde{u} = \frac{1}{e^T\bar{x}}\right)$ is a solution of Mixed LCP (3). On the other hand if $(\tilde{x}, \tilde{u}, \tilde{w})$ is a solution of Mixed LCP (3), then

$$\tilde{u} = \tilde{u}(\tilde{x}^T e) = \tilde{x}^T(A+I)\tilde{x} - \tilde{x}^T\tilde{w} > 0$$

since $(A+I)$ is a strictly copositive matrix. Then $\left(\bar{x} = \frac{\tilde{x}}{\tilde{u}}, \bar{w} = \frac{\tilde{w}}{\tilde{u}}\right)$ is a solution of LCP (1). □

3 Lemke's algorithm for finding an approximate MIS

Since $(A+I)$ is a strictly copositive matrix, then LCP (1) has at least a solution which can be found by Lemke's method [12]. This procedure is a pivotal algorithm that uses basic feasible solutions of the following General LCP ($GLCP$)

$$\begin{array}{l} w = -e + x_0\,d + (A+I)x \\ x \geq 0,\ w \geq 0,\ x_0 \geq 0 \\ x^T w = 0 \end{array} \tag{4}$$

where $d \in \mathbb{R}^n$ is a positive vector (usually $d = e$). To guarantee such a property the algorithm starts with a basic solution of $GLCP$ (4) given by

$$\bar{x} = 0, \quad \bar{x}_0 = \frac{1}{d_r}, \quad w = e + \bar{x}_0 d, \tag{5}$$

where $d_r = \min\{d_i : i = 1,\ldots,n\} > 0$. All the variables x_i and the variable w_r are nonbasic and the remaining variables are basic. Hence there exists exactly a complementary pair (w_r, x_r) of nonbasic variables. In the next iteration the algorithm chooses x_r (complementary of the previous leaving variable) as the entering variable which interchanges with a leaving basic variable (found by the common minimum quotient rule) [5, 16]. A new basic solution of the $GLCP$ (4) is obtained and either $x_0 = 0$ and a solution of the LCP (1) is at hand or the procedure is repeated.

In order to describe the steps of the algorithm, we rewrite the linear constraints of $GLCP(\bar{u})$ in the form

$$Az = q,\, z \geq 0 \qquad (6)$$

where $A = [I \ \ -M \ \ -d\,]$, $z = [\,w \ \ x \ \ x_0\,]^T$ and $q = -e \in \mathbb{R}^n$. We recall [15] that a basic feasible solution for the system (6) is defined by two sets $J = \{k_1, \ldots, k_n\}$ and $L = \{1, \ldots, 2n+1\} \setminus J$ of basic and nonbasic variables respectively, such that

$$z_{k_i} = \bar{q}_i - \sum_{j \in L} \bar{a}_{ij}\, z_j,\ i = 1, \ldots, n \qquad (7)$$

This solution is given by $z_{k_i} = \bar{q}_i,\ i = 1, \ldots, n,\ z_j = 0,\ j \in L$ and is called a basic feasible solution of the $GLCP$ (4). The steps of Lemke's method for solving LCP (1) can now be presented as follows.

LEMKE'S METHOD

Initial Step

Let $d_r = \min\{d_i, i = 1, \ldots, n\}$ and start with a basic feasible solution of $GLCP$ (4), where the x_i variables and w_r are nonbasic and the remaining variables are basic.

General Step

Let $z_s, s \in L$ be the nonbasic variable that is complementary of the variable that has become nonbasic in the previous iteration.

(i) If $\bar{a}_{is} \leq 0$ for all $i = 1, \ldots, n$, stop the algorithm with the termination in an unbounded ray.

(ii) Compute $t \in \{1, \ldots, n\}$ such that

$$t = \min\left\{ r : \frac{\bar{q}_r}{\bar{a}_{rs}} = \min\left\{ \frac{\bar{q}_i}{\bar{a}_{is}} : \bar{a}_{is} > 0 \right\} \right\}. \qquad (8)$$

Perform a pivotal operation which interchanges the nonbasic variable z_s with the basic variable z_{k_t} associated with line t in (8). If $x_0 = 0$ after such an operation, stop the algorithm with a solution of the LCP. Otherwise, repeat the general step.

We suggest [5, 16] for a detailed explanation of Lemke's method, its termination in a finite number of iterations and its implementation for solving medium and large scale $LCPs$. Since $(A+I)$ is strictly copositive [5, 16], then termination in an unbounded ray can not occur [5, 16] and the algorithm terminates with a solution of the LCP.

The choice of the auxiliary vector $d \in \mathbb{R}^n$ has an important impact on the efficiency of the algorithm. We use $d = e$ in order to work with integer basic $GLCP$ solutions. Hence these solutions are highly degenerate and Bland's rule [5, 16] has to be employed for the algorithm to converge. This rule (8) requires an ordering of the rows for its application. In this paper we suggest to apply Lemke's algorithm with Bland's rule n times, where the row orderings are given as follows:

$$(1, 2, \ldots, n{-}1, n),\ (2, \ldots, n{-}1, n, 1),\ \ldots,\ (n, 1, 2, \ldots, n{-}1).$$

Lemke's algorithm with the Bland's rule is able to compute in general different solutions for the LCP (1). It is also interesting to add that the algorithm requires exactly p iterations to find an independent set of cardinal p [13]. So the multistarting Lemke's algorithm finds n independent sets S_i associated to each one of the LCP solutions (some of them may be equal) in $\sum_{i=1}^{n} |S_i|$ iterations, where $|S_i|$ denotes the cardinal of S_i. Since each one of these $|S_i|$ is usually small, then the multistarting algorithm finds these independent sets in a relatively small amount of effort. An approximate independent set of an MIS is given as the one with the largest cardinal.

4 A Sequential complementary algorithm for finding an MIS

There are a number of algorithms for finding a global minimum for the $MPLCC_2$ [2, 4, 7, 8]. A Sequential Complementary (SC) algorithm has been introduced in [11] and searches such a global minimum by solving a parametric $GLCP$, where the parameter strictly reduces in each iteration. By exploiting Theorem 1, it is possible to state the steps of the SC algorithm for solving the $MPLCC_2$ in the following form.

SC ALGORITHM

Initial Step

Find an initial independent set S of G.

General Step

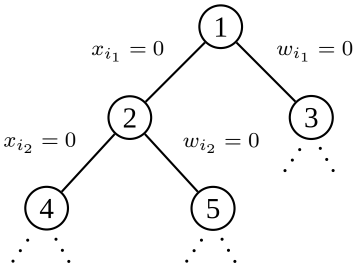
Set $\bar{u} = \dfrac{1}{1+|S|}$ and solve $GLCP(\bar{u})$:

$$\begin{aligned} & w = (A+I)x - ue \\ & e^T x = 1 \\ & x \geq 0, w \geq 0 \\ & u \leq \bar{u} \\ & x^T w = 0. \end{aligned}$$

If this $GLCP(\bar{u})$ has no solution, then S and $|S|$ are an MIS and $\alpha(G)$ respectively. Otherwise, let $\bar{x}$ be a solution of this $GLCP(\bar{u})$. Let $S = \{i : \bar{x}_i > 0\}$ and repeat general step.

It is important to add that the multistarting Lemke's method can be applied to find an initial independent set S. The existence of the constraint $u \leq \bar{u}$ prevents the use of the Lemke's method to solve $GLCP(\bar{u})$. Instead, an enumerative method [1, 10] can be useful in this case.

This algorithm finds a solution of the Mixed LCP (3) by exploring a binary tree of the form



In each node k of the tree, a quadratic program (QP) is considered which consists of minimizing the complementary function $x^T w$ on a convex set defined by the linear constraints of the $GLCP(\bar{u})$ and some constraints $z_i = 0$, where z_i is a complementary x_i or w_i variable that has been fixed to zero in the edges of the path linking the root to the node k. For instance, the QP associated with node 5 of figure above is defined as follows

$$\begin{aligned} QP(5) : Minimize \quad & x^T w \\ subject\ to \quad & w = -e + (A+I)x - ue \\ & e^T x = 1 \\ & x \geq 0, w \geq 0, u \leq \bar{u} \\ & x_{i_1} = w_{i_2} = 0 \end{aligned} \tag{9}$$

A local QP algorithm can be used to find a stationary point $(\tilde{x}, \tilde{w}, \tilde{u})$ for each $QP(t)$. If $\tilde{x}^T \tilde{w} = 0$ then the algorithm stops with a solution of $GLCP(\bar{u})$. Otherwise there must exist at least a complementary pair (x_i, w_i) of positive variables, and two further nodes should be added to the list of open nodes in the tree. The steps of the algorithm are stated below.

ENUMERATIVE ALGORITHM

Step 0

Let $\mathcal{L} = \{1\}$ be the initial list of open nodes.

Step 1

If $\mathcal{L} = \emptyset$, stop: $GLCP(\bar{u})$ has no solution. Otherwise choose a node $t \in \mathcal{L}$.

Step 2

Remove node t from the list $\mathcal{L}$ and consider the quadratic program $QP(t)$ associated with node t. If $QP(t)$ is infeasible, go to Step 1. Otherwise find a stationary point $(\tilde{x}, \tilde{w}, \tilde{u})$ for $QP(t)$.

Step 3

If $\tilde{x}^T \tilde{w} = 0$, stop the algorithm with a solution $(\tilde{x}, \tilde{w}, \tilde{u})$ of $GLCP(\bar{u})$. Otherwise let (x_r, w_r) a pair of complementary positive variables of this stationary point.

Step 4

Add two new nodes k and $k+1$ to the list $\mathcal{L}$ with quadratic programs $QP(k)$ and $QP(k+1)$ defined by
$QP(k) :\ QP(t)$ with constraint $x_r = 0$
$QP(k+1) :\ QP(t)$ with constraint $w_r = 0$
Go to Step 1.

The algorithm possesses finite convergence to a solution of $GLCP(\bar{u})$ or provides a certificate that such a solution does not exist. An implementation of this algorithm can be designed by using the well-known MINOS code [14] to solve all the $QP(t)$ introduced during the whole procedure. Such an implementation also requires efficient heuristic rules for choosing node t in Step 1 and the pair of complementary variables (x_r, w_r) in Step 3. We suggest [10] for a detailed explanation of these heuristic techniques.

It follows from the description of its steps that the enumerative algorithm searches a finite number of stationary points of quadratic programs $QP(t)$ until computing a solution of the $GLCP(\bar{u})$ or showing that no solution exists. The zero value of the complementary function $x^T w$ is a stopping criterion when $GLCP(\bar{u})$ has a solution. However, this criterion is never satisfied when $GLCP(\bar{u})$ has no solution and the algorithm requires much more effort to terminate in Step 1 with an empty list $\mathcal{L}$. So, the SC algorithm incorporating this enumerative method can in general find a global minimum of an $MPLCC$ in a reasonable amount of effort but may face too many difficulties to give a certificate that such a global minimum has been achieved.

5 Computational experience

The test problems used in our experiences are graphs from the Dimacs collection [9]. The number of nodes n, edges $|E|$ and the independent number $\alpha(G)$ of these graphs are included in Table 1. For each value of n, $|E|$ represents the number of edges of the complement of the graph with n nodes given in the Dimacs data set. Table 2 reports the results of the experiences for finding $\alpha(G)$ by the complementary approaches discussed in the previous sections on a Pentium IV 2.4 GHz machine with 512 MB of RAM.

The notations InLem and InSc represent the cardinal of the independent sets found by the multistarting Lemke's method and the SC algorithm with the previous method in the initial step respectively. Furthermore Nis and Nit denote the number of pivotal operations required by the SC algorithm to find an MIS (Nis$= 0$ when the multistarting Lemke's method provides such a set) and to terminate the algorithm (show that the last $GLCP(\bar{u})$ has no solution) respectively. Finally T is the total CPU time in seconds spent by the SC algorithm.

Table 1:

Problem	n	$\lvert E\rvert$	$\alpha(G)$
matrixj1	28	168	4
matrixj2	70	560	14
matrixj3	120	1680	8
matrixj4	496	14880	16
matrixm1	45	72	16
matrixm2	378	702	126
matrixh1	64	192	32
matrixh2	64	1312	4
matrixh3	256	1024	128
matrixh4	256	11776	16
matrixb1	200	5066	21
matrixk4	171	5100	11
matrixf21	200	18366	12
matrixf22	200	16665	24
matrixf25	200	11427	58
matrixp2	300	22922	25
matrixsr27	200	6032	18
matrixsr29	200	2037	42

Table 2:

Problem	InLem	InSc	Nis	Nit	T
matrixj1	4	4	0	4812	1.02
matrixj2	14	14	0	*	13.70
matrixj3	8	8	0	*	19.95
matrixj4	16	16	0	*	71.15
matrixm1	16	16	0	*	19.85
matrixm2	125	126	33249	*	173.26
matrixh1	32	32	0	*	10.46
matrixh2	4	4	0	17828	6.29
matrixh3	128	128	0	*	79.18
matrixh4	16	16	0	*	44.41
matrixb1	18	21	3518	*	66.70
matrixk4	9	11	221	*	45.29
matrixf21	12	12	0	*	102.31
matrixf22	24	24	0	*	108.15
matrixf25	58	58	0	*	73.21
matrixp2	22	25	261	*	80.90
matrixsr27	16	18	558	*	40.22
matrixsr29	35	42	9019	*	38.64

The numerical results reported in Table 2 show that the multistarting Lemke's method gives in general a very good approximate for $\alpha(G)$. The algorithm is even able to find $\alpha(G)$ in many cases. Furthermore the SC algorithm with the multistarting Lemke's method in its initial step has computed $\alpha(G)$ for all the test problems in a reasonable amount of effort. However, the algorithm is in general unable to give a certificate that an MIS has been achieved in less than 50000 pivotal operations (an asterisk is included in Table 2 for these cases). We believe that the enumerative algorithm employed by the SC method can be improved not only to reduce the computational effort to find an MIS but also to provide the required certificate. This is a topic of our current research.

References:

[1] F. Al-Khayyal, An implicit enumeration procedure for the general linear complementarity problem, *Mathematical Programming Studies*, Vol.31, 1987, pp. 1–20.

[2] J. Bard and J. Moore, A branch and bound algorithm for the bilevel programming problem, *SIAM Journal on Scientific and Statistical Computing*, Vol.11(2), 1990, pp. 281–292.

[3] I. M. Bomze, On standard quadratic optimization problems, *Journal of Global Optimization*, Vol.13, 1998, pp. 369–387.

[4] S. Burer and D. Vandenbussche, A finite branch-and-bound algorithm for nonconvex quadratic programming via semidefinite relaxations, *Mathematical Programming, Series A*, Vol.113, 2008, pp. 259–282.

[5] R. Cottle, J. Pang and R. Stone, The Linear Complementarity Problem, Academic Press, New York, 1992.

[6] D. Ding-Zhu and P. M. Pardalos, Handbook of Combinatorial Optimization, Kluwer Academic Publishers, Dordrecht, 1999.

[7] P. Hansen, B. Jaumard and G. Savard, New branch-and-bound rules for linear bilevel programming, *SIAM Journal on Science and Statistical Computing*, Vol.13(5), 1992, pp. 1194–1217.

[8] J. Hu, J. Mitchell, J. Pang, K. Bennett and G. Kunapuli, On the global solution of linear programs with linear complementarity constraints, *SIAM Journal on Optimization*, Vol.19(1), 2008, pp. 445–471.

[9] D. Johnson and A. Trick, Cliques, Coloring and Satisfiability: Second DIMACS Implementation Challenge, *American Mathematical Society*, 26th of Dimacs series, 1996.

[10] J. Júdice and A. Faustino, An experimental investigation of enumerative methods for the linear complementarity problem, *Computers and Operations Research*, Vol.15, 1988, pp. 417–426.

[11] J. Júdice and A. Faustino, A SLCP method for bilevel linear programming, *Annals of Operations Research*, Vol.34, 1992, pp. 89–106.

[12] C. Lemke, On Complementary pivot theory, *Mathematics of Decision Sciences*, G. Dantzig and A. Veinott Jr. editors, American Mathematical Society Providence, 1968, pp. 95–114.

[13] A. Massaro, M. Pelillo and I. Bomze, A complementary pivoting approach to the maximum weight clique problem, *SIAM Journal on Optimization*, Vol.12(4), 2002, pp. 928–948.

[14] B. Murtagh and M. Saunders, MINOS 5.1, User's Guide, Stanford University, 1987.

[15] K. Murty, Linear Programming, John Wiley & Sons, New York, 1983.

[16] K. Murty, Linear Complementarity, Linear and Nonlinear Programming, Heldermann Verlag, Berlin, 1988.