RAIRO Operations Research

RAIRO Oper. Res. **37** (2003) 17-27 DOI: 10.1051/ro:2003012

A DERIVATION OF LOVÁSZ' THETA VIA AUGMENTED LAGRANGE DUALITY

Mustapha Ç. Pinar¹

Communicated by Jean Abadie

Abstract. A recently introduced dualization technique for binary linear programs with equality constraints, essentially due to Poljak *et al.* [13], and further developed in Lemaréchal and Oustry [9], leads to simple alternative derivations of well-known, important relaxations to two well-known problems of discrete optimization: the maximum stable set problem and the maximum vertex cover problem. The resulting relaxation is easily transformed to the well-known Lovász θ number.

Keywords. Lagrange duality, stable set, Lovász theta function, semidefinite relaxation.

Mathematics Subject Classification. 90C27, 90C27, 90C35.

1. Background

The problem of finding a stable set or equivalently, a maximal (or, a maximal weighted independent set in weighted graphs) independent set in a graph is one of the most difficult problems of combinatorial optimization. It is known to be NP-complete for arbitrary graphs. Furthermore, it is also very difficult to approximate as mentioned in [6, 14].

Lovász was the first one to introduce upper bounds of semidefinite type for this problem where his investigations were motivated by some problems in information theory [10,11]. In particular, Shannon, studying the problems of interference stable

¹ Bilkent University, Department of Industrial Engineering, 06800 Ankara, Turkey; e-mail: mustafap@bilkent.edu.tr.

Received July, 2002.

coding in 1956, introduced the concept of information capacity of a graph which is intimately related to the maximal independent set of the graph. However, Shannon's measure for the information capacity of a graph turns out to be a function which is very hard to compute even for simple graphs like the pentagon circuit C_5 (a circuit with five nodes). It was only in 1977 that Lovász obtained the precise result that the Shannon capacity of C_5 was equal to $\sqrt{5}$. In fact, Lovász's result gave a polynomially computable upper bound (by a judicious use of the ellipsoid method) on the maximal independent set of an arbitrary graph. Furthermore, for a class of graphs called *perfect* graphs, Lovász's bound is exact. This bound is commonly referred to as Lovász theta function (or, number). Lovász theta function can be computed as the solution of a semidefinite programming problem (there exist several formulations; see *e.g.* [3,7]), which spawned a flurry of activities in the numerical optimization community with the advent of polynomial interior point methods in the 80's.

1.1. Augmented Lagrange duality and semidefinite relaxations

Poljak, Rendl and Wolkowicz proposed a novel dualization technique which they called a recipe for obtaining semidefinite programming relaxations for quadratic 0-1 programs in [13] using redundant constraints in an augmented Lagrangian framework. In a recent paper [9], Lemaréchal and Oustry extended the technique to linear integer programs with equality constraints, which results in a semidefinite programming relaxation of the problem. In summary, the idea is the following. Consider the linear integer programming problem

maximize
$$c^T x$$

s.t. $Ax = b$
 $x_i \in \{0, 1\}, \forall i \in \{1, \dots, n\}$.

Lemaréchal and Oustry rewrite this problem by adding a redundant quadratic constraint and treating the 0 - 1 constraints as quadratic equations as follows:

maximize
$$c^T x$$

s.t. $Ax = b$
 $\|Ax - b\|_2^2 = 0$
 $x_i^2 = x_i, \forall i \in \{1, \dots, n\}$.

Then, they form a Lagrange dual of the above problem, and take the dual of the resulting problem one more time to arrive at the following convex, semidefinite programming relaxation (bi-dual) of the original problem:

maximize
$$c^T d(X)$$

s.t. $Ad(X) = b$
Trace $A^T A X = ||b||_2^2$
 $\begin{bmatrix} 1 & d(X)^T \\ d(X) & X \end{bmatrix} \succeq 0$

with X a symmetric $n \times n$ matrix, and d(X) the vector composed of its diagonal elements. Appending the redundant constraint with a scalar multiplier to the Lagrange function is reminiscent of augmented Lagrangian methods, hence our title. Lemaréchal and Oustry then establish that this technique is equivalent to what they propose in their paper as "Dualization B", which essentially consists in dualizing the linear inequality constraints, and minimizing the resulting Lagrange function over the quadratic constraints resulting from the binary nature of the variables. Under certain technical conditions, they show that maximizing the resulting function ("Dualization B") gives a better bound than the linear programming relaxation of the original problem.

Our purpose in this note is to explore the consequences of this dualization technique in the context of the maximum (weighted) stable set problem and the maximum (weighted) vertex cover problem. Given the importance of these problems both from theoretical and applied viewpoints, this note adds to the repertoire of many derivations of Lovász's theta (see *e.g.* [7] for a detailed exposition of these derivations) yet another simple and concise derivation.

2. The maximum (weighted) stable set problem

We consider the linear integer programming formulation of the stable set problem on a connected graph G = (V, E) (with node set V and edge set E), referred to as (SSP):

maximize
$$c^T x$$

s.t. $x_i + x_j \le 1, \forall (i, j) \in E$
 $x_i \in \{0, 1\}, \forall i \in V$

where $c \in \mathbf{R}^{|V|}$ is a positive vector. The optimal value is referred to as the (weighted) independence number, $\alpha(G)$, of G. The problem is also sometimes called the node packing or vertex packing problem.

To view the stable set problem in the context of Lemaréchal–Oustry we reformulate the problem as follows:

maximize
$$c^T x$$

s.t. $x_i + x_j + s_{ij} = 1, \forall (i, j) \in E$
 $x_i \in \{0, 1\}, \forall i \in V$
 $s_{ij} \in \{0, 1\}, \forall (i, j) \in E$

and, treating the binary constraints as quadratic constraints

maximize
$$c^T x$$

s.t. $x_i + x_j + s_{ij} = 1, \forall (i, j) \in E$
 $x_i^2 = x_i, \forall i \in V$
 $s_{ij}^2 = s_{ij}, \forall (i, j) \in E$

we can apply the dualization B technique. We obtain the following SDP problem which we refer to (SSLO):

maximize
$$c^T d(X)$$

s.t. $Pd(X) + d(S) = e$
Trace $A^T A \mathbf{X} = |E|$
 $\begin{bmatrix} 1 & d(\mathbf{X})^T \\ d(\mathbf{X}) & \mathbf{X} \end{bmatrix} \succeq 0$

where P represents the $|E| \times |V|$ edge-node incidence matrix, A represents the block matrix $A = \begin{bmatrix} P & I \end{bmatrix}$ with

$$A^T A = \left[\begin{array}{cc} P^T P & P^T \\ P & I \end{array} \right],$$

the (variable) symmetric matrix \mathbf{X} is as follows:

$$\mathbf{X} = \left[\begin{array}{cc} X & B^T \\ B & S \end{array} \right],$$

e is a vector of all ones, and, finally $d(X) \in \mathbf{R}^{|V|}$ denotes the diagonal of X, $d(S) \in \mathbf{R}^{|E|}$ the diagonal of S, respectively.

The intriguing question here is whether this is a new relaxation to the stable set problem. To answer this question, we have to relate the value of the above relaxation which we denote z_{sslo} to well-known relaxations of the stable set problem. The most famous relaxation of the stable set problem is the θ -number or θ -function of Lovász, for which at least seven different formulations are known; see [3,7]. Can we transform the above relaxation into one of the equivalent forms of Lovász θ ?

Let us begin by inspecting closely the constraints of the relaxation (SSLO). Obviously, the first set of constraints are nothing else than the constraints

$$x_i + x_j + s_{ij} = 1, \ \forall (i,j) \in E$$

of the linear integer programming formulation, expressed using the diagonal elements of the large matrix **X**. The second constraint of (SSLO) has three components, (1) the component Trace $P^T P X$ which expresses the connectivity properties of the graph, (2) 2 * Trace $P^T B$, and (3) the term Trace S which is just the sum of the diagonal elements of S.

Now, consider dropping the matrix B altogether, and the off-diagonal elements of S which do not seem to play any role in the problem. In fact, we can reduce Sto vector s, which is just its diagonal. This leaves us with the SDP problem we refer to as SSSLO (simplified SSLO):

maximize
$$c^T d(X)$$

s.t. $Pd(X) + s = e$
Trace $P^T P X + e^T s = |E|$
 $\begin{bmatrix} 1 & d(X)^T \\ d(X) & X \end{bmatrix} \succeq 0$
 $s \ge 0.$

It is easy to verify that SSSLO is still a valid relaxation of the stable set problem. Just take a stable set in the graph G and its incidence vector x, and form the matrix $X = xx^T$. This is a feasible solution to SSSLO along with the accompanying slack vector. It is also immediate that $z_{ssslo} \leq z_{sslo}$ since SSSLO is a restriction of SSLO.

Now, a careful look at the second constraint reveals the following structure

$$\sum_{i \in V} \delta_i X_{ii} + \sum_{(i,j) \in E} s_{ij} + 2 \sum_{(i,j) \in E, i \neq j} X_{ij} = |E|$$

where δ_i is the number of nodes adjacent to node *i*. It is obvious using the first set of constraints that the above constraint is simply

$$\sum_{(i,j)\in E, i\neq j} X_{ij} = 0$$

Hence, our relaxation is in fact

maximize
$$c^T d(X)$$

s.t. $Pd(X) + s = e$
 $\sum_{\substack{(i,j) \in E, i \neq j \\ d(X)^T \\ d(X) X}} X_{ij} = 0$
 $\begin{bmatrix} 1 & d(X)^T \\ d(X) & X \end{bmatrix} \succeq 0$
 $s \ge 0.$

Now, let us append to this relaxation the following constraints:

$$X_{ij} \ge 0, \ \forall \ (i,j) \in E.$$

$$\tag{1}$$

With these non-negativity constraints we still conserve the property that the resulting SDP problem is a relaxation of SSP, and that the resulting optimal value, $z_{nnssslo}$, say, is at most as large as z_{ssslo} , *i.e.*, $z_{nnssslo} \leq z_{ssslo}$.

Hence, we have so far looked into three SDP relaxations for SSP with respective optimal values in the following order:

$$z_{nnssslo} \le z_{ssslo} \le z_{sslo}.$$

But, the non-negativity constraints (1) along with the second constraint, namely,

$$\sum_{(i,j)\in E, i\neq j} X_{ij} = 0,$$

imply that $X_{ij} = 0 \forall (i, j) \in E, i \neq j$. Hence, we obtain the following problem:

maximize
$$c^T d(X)$$

s.t. $Pd(X) + s = e$
 $X_{ij} = 0, \forall (i, j) \in E,$
 $\begin{bmatrix} 1 & d(X)^T \\ d(X) & X \end{bmatrix} \succeq 0$
 $s \ge 0.$

On the other hand, the first set of constraints are now redundant. To see this let q_{ij} be a $\mathbf{R}^{|V|+1}$ vector with a -1 in the zeroth position, and 1 in the *i* and *j* positions, respectively. Now, using the fact that $q_{ij}^T \begin{bmatrix} 1 & d(X)^T \\ d(X) & X \end{bmatrix} q_{ij} \ge 0$, and that $X_{ij} = 0 \forall (i, j) \in E, i \neq j$, we obtain the inequality

$$X_{ii} + X_{jj} \le 1, \ \forall (i,j) \in E.$$

$$\tag{2}$$

Therefore, we arrived at the following SDP formulation:

maximize
$$c^T d(X)$$

s.t. $X_{ij} = 0, \forall (i, j) \in E$
 $\begin{bmatrix} 1 & d(X)^T \\ d(X) & X \end{bmatrix} \succeq 0$

which is one of the several formulations of the θ -number of Lovász; see Lemma 2.17 of Lovász and Schrijver [5, 12]. Lemaréchal and Oustry [9] re-derive this form of the θ function by taking the Lagrange dual of the following quadratic formulation of the stable set problem:

maximize
$$c^T x$$

s.t. $x_i x_j = 0, \forall (i, j) \in E$
 $x_i \in \{0, 1\}, \forall i \in V,$

and taking the dual of the resulting semidefinite program.

2.1. Observations and discussion

We have the following observations.

(1) It is not true that the constraints defining (SSSLO), namely,

$$X_{ii} + X_{jj} \le 1, \ (i, j) \in E,$$
$$\sum_{(i, j) \in E} X_{ij} = 0$$
$$\begin{bmatrix} 1 & d(X)^T \\ d(X) & X \end{bmatrix} \succeq 0$$

imply the constraints:

$$X_{ij} \ge 0, \ (i,j) \in E.$$

As an example consider the graph K_3 and the matrix

$$Z := \begin{pmatrix} 1 & x & x & x \\ x & x & y & -y \\ x & y & x & 0 \\ x & -y & 0 & x \end{pmatrix}$$

Take for instance, x = 1/4 and y = 1/8. Then, Z is feasible for (SSSLO).

- (2) For complete graphs, it is easy to see that $z_{ssslo} = \theta$. When c = e, it is well-known that $\theta = 1$, and $z_{ssslo} \leq 1$ can be seen from the inequality $f^T Y f \geq 0$ for f = (-1, 1, ..., 1) and $Y = \begin{bmatrix} 1 & d(X)^T \\ d(X) & X \end{bmatrix}$.
- (3) We have conducted numerical experiments using the semidefinite programming software packages SDPHA [2] and SDPPACK [1]. In particular, we have solved the problem (SSSLO) using SDPHA and computed Lovász theta for the same graph using a built-in function in SDPPACK. In all our experiments, including odd circuits with up to 11 nodes (with c = e), and other small examples with weighted graphs or unit costs, we always observed equality between z_{ssslo} and theta. Notice that it is elementary to see that $\theta \leq z_{ssslo}$. Further research is required to establish or refute this claim of equality between the two numbers.
- (4) In reference to 1 above, we always obtained an optimal matrix X with $X_{ij} = 0$ for $(i, j) \in E$ in our numerical experiments.
- (5) Another interesting observation based on our computational experience seems to suggest that the matrix X corresponding to an optimal solution to (SSSLO) in the case of odd circuit graphs has a circulant structure, e.g., for C_5 , SDPHA reports the following optimal matrix

$$Z := \begin{pmatrix} 1 & 0.4472 & 0.4472 & 0.4472 & 0.4472 & 0.4472 \\ 0.4472 & 0.4472 & 0 & 0.2764 & 0.2764 & 0 \\ 0.4472 & 0 & 0.4472 & 0 & 0.2764 & 0.2764 \\ 0.4472 & 0.2764 & 0 & 0.4472 & 0 & 0.2764 \\ 0.4472 & 0.2764 & 0.2764 & 0 & 0.4472 & 0 \\ 0.4472 & 0 & 0.2764 & 0.2764 & 0 & 0.4472 \end{pmatrix}$$

In addition to the circulant structure, for C_3, C_5, C_7, C_9 and C_{11} (with c = e), we obtain the optimal diagonal values as $X_{ii}(2k+1) = \frac{\cos \frac{\pi}{2k+1}}{1+\cos \frac{\pi}{2k+1}}$ for $i = 1, \ldots, 2k+1$ (Lovász [11] had proved that $\theta(C_{2k+1}) = \frac{(2k+1)\cos \frac{\pi}{2k+1}}{1+\cos \frac{\pi}{2k+1}}$).

3. The minimum (weighted) cover problem

We consider now the linear integer programming formulation of the minimum vertex cover problem on a connected graph G = (V, E) (VCP):

minimize
$$c^T x$$

s.t. $x_i + x_j \ge 1, \forall (i, j) \in E$
 $x_i \in \{0, 1\}, \forall i \in V$

where $c \in \mathbf{R}^{|V|}$ is a positive vector. It is well-known that the value of the minimum vertex cover, vc(G), say, is related to the independence number $\alpha(G)$ as

$$vc(G) + \alpha(G) = W \tag{3}$$

where $W = \sum_{i \in V} c_i$, (or, as $vc(G) + \alpha(G) = |V|$).

Adding binary surplus variables, treating the binary constraints as quadratic constraints we obtain

minimize
$$c^T x$$

s.t. $x_i + x_j - s_{ij} = 1, \forall (i, j) \in E$
 $x_i^2 = x_i, \forall i \in V$
 $s_{ij}^2 = s_{ij}, \forall (i, j) \in E.$

The augmented Lagrange duality technique yields the following relaxation as bidual (VCR1):

minimize
$$c^T d(X)$$

s.t. $Pd(X) - d(S) = e$
Trace $A^T A \mathbf{X} = |E|$
 $\begin{bmatrix} 1 & d(\mathbf{X})^T \\ d(\mathbf{X}) & \mathbf{X} \end{bmatrix} \succeq 0$

where P represents the $|E|\times |V|$ edge-node incidence matrix, A represents the block matrix $A=\left[\begin{array}{cc}P&-I\end{array}\right]$ with

$$A^T A = \left[\begin{array}{cc} P^T P & -P^T \\ -P & I \end{array} \right],$$

the (variable) symmetric matrix \mathbf{X} is as follows:

$$\mathbf{X} = \left[\begin{array}{cc} X & B^T \\ B & S \end{array} \right] \cdot$$

Denote its optimal value z_{vcr1} . Using arguments similar to those of the previous section, we simplify this relaxation to the following semidefinite program (VCLO).

minimize
$$c^T d(X)$$

s.t. $Pd(X) - s = e$
 $\sum_{\substack{(i,j) \in E, i \neq j \\ d(X) = 0}} (X_{ij} - s_{ij}) = 0$
 $\begin{bmatrix} 1 & d(X)^T \\ d(X) = X \end{bmatrix} \succeq 0$
 $s \ge 0.$

It is immediate to verify that VCLO is a relaxation of VCP. Just take a minimum vertex cover in the graph G and its incidence vector x, and form the matrix $X = xx^{T}$. This is a feasible solution to VCLO along with the accompanying slack vector. Now, append to this relaxation the following constraints:

$$X_{ij} \ge s_{ij} \forall (i,j) \in E.$$

These imply, together with the second set of constraints that,

$$X_{ij} = s_{ij} \ \forall \ (i,j) \in E,$$

or, equivalently,

$$X_{ij} = X_{ii} + X_{jj} - 1 \ \forall \ (i,j) \in E$$

Therefore, we have arrived at the relaxation

minimize
$$c^T d(X)$$

s.t. $Pd(X) \ge e$
 $X_{ij} - X_{ii} - X_{jj} = -1, \forall (i, j) \in E,$
 $\begin{bmatrix} 1 & d(X)^T \\ d(X) & X \end{bmatrix} \succeq 0.$

As is the case with the stable set relaxation, the first set of constraints are now redundant. Therefore, we reach the relaxation (VCSDP):

minimize
$$c^T d(X)$$

s.t. $X_{ij} - X_{ii} - X_{jj} = -1, \forall (i, j) \in E,$
 $\begin{bmatrix} 1 & d(X)^T \\ d(X) & X \end{bmatrix} \succeq 0.$

Denote its optimal value by z_{vcsdp} .

Now, it is well-known, e.g. [6, 8], that a similar relation to 3 holds between the respective semidefinite relaxations of minimum vertex cover problem and the stable set problem, namely:

$$sdp(G) + \theta(G) = W \tag{4}$$

where sdp(G) is the optimal value of the following program (VCMC):

minimize
$$\sum_{i \in V} c_i \frac{1 + Y_{0i}}{2}$$

s.t.
$$Y_{ij} - Y_{0i} - Y_{0j} = -1, \forall (i, j) \in E,$$
$$d(Y) = e$$
$$Y \succeq 0.$$

Notice that we can also obtain this relaxation departing from the quadratic programming formulation of minimum vertex cover, namely,

minimize
$$c^T x$$

s.t. $(1-x_i)(1-x_j) = 0, \forall (i,j) \in E$
 $x_i \in \{0,1\}, \forall i \in V,$

using the same steps as Lemaréchal and Oustry [9].

On the other hand, VCMC is equivalent to VCSDP via the bijective transformation $\tilde{X} = QYQ^T$ where

$$\tilde{X} = \left[\begin{array}{cc} 1 & d(X)^T \\ d(X) & X \end{array} \right]$$

and the $(n+1) \times (n+1)$ matrix Q is given by

$$Q = \left[\begin{array}{cc} 1 & 0\\ \frac{1}{2}e & \frac{1}{2}I_n \end{array} \right],$$

by Proposition 5.2 of Helmberg [4]. *I.e.*, the mapping $\phi : S_{n+1} \mapsto S_{n+1}, Y \mapsto X = QYQ^T$ where S_{n+1} denotes the space of $(n + 1) \times (n + 1)$ symmetric matrices, bijectively maps feasible solutions of VCMC to VCSDP, and with equal objective function values. Therefore, we have that

$$z_{vcsdp} + \theta(G) = W.$$

As a final remark, we observed as in the stable set case, through computational experiments that there is already equality between z_{vcr1} , z_{vclo} and $W - \theta(G)$.

The author would like to thank Francesco Maffioli for suggesting that the author prepare this note.

References

- F. Alizadeh, J.-P. Haberly, V. Nayakkankuppam and M.L. Overton, SDPPACK user's guide, Technical Report 734. NYU Computer Science Department (1997).
- [2] N. Brixius, R. Sheng and F. Potra, SDPHA user's guide, Technical Report. University of Iowa (1998).

- M. Grötschel, L. Lovász and A. Schrijver, Geometric Algorithms and Combinatorial Optimization. Springer-Verlag, Berlin (1988).
- [4] C. Helmberg, Fixing variables in semidefinite relaxations. SIAM J. Matrix Anal. Appl. 21 (2000) 952-969.
- [5] C. Helmberg, S. Poljak, F. Rendl and H. Wolkowicz, Combining semidefinite and polyhedral relaxations for integer programs, edited by E. Balas and J. Clausen, Integer Programming and Combinatorial Optimization IV. Springer-Verlag, Berlin, Lecture Notes in Comput. Sci. 920 (1995) 124-134.
- [6] J. Kleinberg and M. Goemans, The Lovász theta function and a semidefinite programming relaxation of vertex cover. SIAM J. Discrete Math. 11 (1998) 196-204.
- [7] D. Knuth, The sandwich theorem. *Electron. J. Combinatorics* 1 (1994);
 www.combinatorics.org/Volume1.html#A1
- [8] M. Laurent, S. Poljak and F. Rendl, Connections between semidefinite relaxations of the max-cut and stable set problems. *Math. Programming* 77 (1997) 225-246.
- C. Lemaréchal and F. Oustry, Semidefinite relaxation and Lagrangian duality with application to combinatorial optimization, Technical Report 3170. INRIA Rhône-Alpes (1999); http://www.inria.fr/RRRT/RR-3710.html
- [10] L. Lovász, On the Shannon capacity of a graph. IEEE Trans. Inform. Theory 25 (1979) 355-381.
- [11] L. Lovász, Bounding the independence number of a graph. Ann. Discrete Math. 16 (1982).
- [12] L. Lovász and L. Schrijver, Cones of matrices, and set functions and 0 1 optimization. SIAM J. Optim. 1 (1991) 166-190.
- [13] S. Poljak, F. Rendl and H. Wolkowicz, A recipe for semidefinite relaxation for (0 1) quadratic programming. J. Global Optim. 7 (1995) 51-73.
- [14] N. Shor, Nondifferentiable Optimization and Polynomial Problems. Kluwer Academic Publishers, Dordrecht, The Netherlands (1998).

To access this journal online: www.edpsciences.org