9 The stable set problem and the Lovász theta function

We now look at another application of semidefinite optimisation to combinatorial optimisation, namely to the maximum stable set problem.

Stable set Let $G = (V, E)$ be an undirected graph. A *stable set* (also known as an *independent set*) in G is a subset $S \subseteq V$ such that no two vertices in S are connected by an edge, i.e., $i, j \in S \Rightarrow \{i, j\} \notin E$. The *maximum stable set problem* is the problem of finding the largest stable set in a graph. The stable set problem can be formulated as the following problem:

$$\begin{array}{lll} \underset{x\in\mathbb{R}^n, X\in\mathbf{S}^n}{\text{maximise}} & \displaystyle\sum_{i=1}^n x_i & \\ \text{subject to} & x_i^2 = x_i & \forall i \in V = \{1,\dots,n\} \\ & x_i x_j = 0 & \forall ij \in E. \end{array} \tag{1}$$

The constraint $x_i^2 = x_i$ is equivalent to saying that $x_i \in \{0, 1\}$ and the stable set S corresponds to the set of i such that $x_i = 1$. Note that the constraint $x_i x_j = 0$ ensures that S is a stable set. The objective function $\sum_{i=1}^n x_i$ counts the cardinality of S. Solving the optimisation problem (1) is computationally hard in general.

Semidefinite relaxation We are now going to define a semidefinite relaxation for (1). This relaxation was first proposed by Lovász in [Lov79]. It allows us to get an upper bound on the solution (1) by solving a semidefinite program.

$$\begin{array}{lll} \underset{x\in\mathbb{R}^n, X\in\mathbf{S}^n}{\text{maximise}} & \displaystyle\sum_{i=1}^n x_i & \\ \text{subject to} & X_{ii} = x_i & i \in V \\ & X_{ij} = 0 & ij \in E \\ & \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \succeq 0 & \end{array} \tag{2}$$

Problem (2) can be solved efficiently using algorithms for semidefinite programming. The next theorem shows that (2) yields an upper bound on (1).

Theorem 9.1. *Let $\alpha(G)$ be the solution of (1) and $\vartheta(G)$ be the solution of (2). Then $\alpha(G) \le \vartheta(G)$.*

Proof. It suffices to observe that if x is feasible for (1), then the pair $(x, X = xx^T)$ is feasible for (2) since

$$\begin{bmatrix} 1 & x^T \\ x & xx^T \end{bmatrix} = \begin{bmatrix} 1 \\ x \end{bmatrix} \begin{bmatrix} 1 \\ x \end{bmatrix}^T \succeq 0.$$

□

A natural question is to ask whether there is a constant $c > 0$ such that $c \cdot \vartheta(G) \le \alpha(G)$ for all graphs G. Unfortunately this is not the case. Indeed one can show:

Theorem 9.2. *There exists a sequence of graphs (G_n) such that $\vartheta(G_n)/\alpha(G_n) \geq \frac{\sqrt{n}}{3\log n} \to \infty$ as $n \to \infty$.*

The proof appears in [BTN01] and relies on the following lemmas. For any graph $G = (V, E)$ we define the *complement graph* of G as $\bar{G} = (V, \bar{E})$ where $\bar{E} = \{\{i, j\} : i \neq j \text{ and } \{i, j\} \notin E\}$. The first lemma gives a semidefinite formulation of $\vartheta(G)$ as a minimization problem.

Lemma 1. *For any graph G we have*

$$\begin{array}{rll} \vartheta(G) = & \textit{min.} & Z_{00} \\ & \textit{s.t.} & Z_{ii} = 1 \qquad \forall i \in V \\ & & Z_{ij} = 0 \qquad \forall ij \in \bar{E} \\ & & \begin{bmatrix} Z_{00} & \mathbf{1}^T \\ \mathbf{1} & Z \end{bmatrix} \succeq 0 \end{array} \tag{3}$$

Proof. We use duality. If we let $\lambda_i \in \mathbb{R}$ be the dual variable for the constraint "$X_{ii} = x_i$" in (2), $\mu_{ij} \in \mathbb{R}$ for the constraints "$X_{ij} = 0$ for $ij \in E$" and $\Gamma = \left[\begin{smallmatrix} Z_{00} & z^T \\ z & Z \end{smallmatrix}\right] \succeq 0$ the dual variable for the SDP constraint in (2) we get that for any feasible (x, X) of (2):

$$\sum_{i \in V} \lambda_i (X_{ii} - x_i) + \sum_{ij \in E} \mu_{ij} X_{ij} + \langle \left[\begin{smallmatrix} Z_{00} & z^T \\ z & Z \end{smallmatrix}\right], \left[\begin{smallmatrix} 1 & x^T \\ x & X \end{smallmatrix}\right] \rangle \geq 0.$$

Rearranging, this is equivalent to

$$\sum_{i \in V} (\lambda_i - 2z_i) x_i - \langle \operatorname{diag}(\lambda) + M + Z, X \rangle \leq Z_{00}.$$

where $M \in \mathbf{S}^n$ is defined as $M_{ij} = \mu_{ij}$ if $ij \in E$ and 0 otherwise. If the dual variables (λ, M, Z) satisfy $-\lambda_i + 2z_i = 1$ for all i and $\operatorname{diag}(\lambda) + M + Z = 0$ and $\left[\begin{smallmatrix} Z_{00} & z^T \\ z & Z \end{smallmatrix}\right] \succeq 0$, then Z_{00} is an upper bound to $\vartheta(G)$. The dual consists of finding the best such upper bound to $\vartheta(G)$ and so can be expressed as:

$$\begin{array}{ll} \underset{\lambda, M, \left[\begin{smallmatrix} Z_{00} & z^T \\ z & Z \end{smallmatrix}\right]}{\text{min.}} & Z_{00} \\ \text{s.t.} & \begin{bmatrix} Z_{00} & z^T \\ z & Z \end{bmatrix} \succeq 0 \\ & \operatorname{diag}(\lambda) + M + Z = 0 \\ & M_{ij} = 0 \text{ if } ij \in \bar{E} \text{ or } i = j \\ & \lambda_i - 2z_i = 1 \qquad \forall i \in V. \end{array}$$

The variables λ and M can be eliminated from the problem above. Indeed the second and third constraints together can be equivalently written as: $Z_{ij} = 0$ for $ij \in \bar{E}$ and $Z_{ii} = -\lambda_i$ for all $i \in V$. If we do this simplification we get:

$$\begin{array}{ll} \text{min.} & Z_{00} \\ \text{s.t.} & \begin{bmatrix} Z_{00} & z^T \\ z & Z \end{bmatrix} \succeq 0 \\ & Z_{ij} = 0 \qquad \forall ij \in \bar{E} \\ & z_i = -(1 + Z_{ii})/2 \qquad \forall i \in V. \end{array} \tag{4}$$

Now using the fact that $\left[\begin{smallmatrix} A & B^T \\ B & C \end{smallmatrix}\right] \succeq 0 \iff \left[\begin{smallmatrix} A & -B^T \\ -B & C \end{smallmatrix}\right] \succeq 0$ this dual problem can be further simplified by turning the minus sign in the last constraint to a plus sign:

$$\begin{array}{ll} \text{min.} & Z_{00} \\ \text{s.t.} & \begin{bmatrix} Z_{00} & z^T \\ z & Z \end{bmatrix} \succeq 0 \\ & Z_{ij} = 0 \qquad \forall ij \in \bar{E} \\ & z_i = (1 + Z_{ii})/2 \qquad \forall i \in V. \end{array} \tag{5}$$

We show that this dual problem can be further simplified to (3). It is clear that the optimal value of (5) is $\leq$ the optimal value of (3) ((3) corresponds to taking $Z_{ii} = 1$ and $z_i = -1$). We need to show the other inequality. Let (z, Z) be feasible points of (5). Define

$$\tilde{Z}_{ij} = \begin{cases} Z_{ij}/z_i z_j & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases}$$

If we show that $\left[\begin{smallmatrix} Z_{00} & \mathbf{1}^T \\ \mathbf{1} & \tilde{Z} \end{smallmatrix}\right] \succeq 0$ we are done. Let $Z'_{ij} = Z_{ij}/(z_i z_j)$. Note that

$$\begin{bmatrix} Z_{00} & \mathbf{1}^T \\ \mathbf{1} & Z' \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \operatorname{diag}(z) \end{bmatrix}^{-1} \begin{bmatrix} Z_{00} & z^T \\ z & Z \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \operatorname{diag}(z) \end{bmatrix}^{-1} \succeq 0.$$

Now note that $Z'_{ii} = Z_{ii}/((1 + Z_{ii})/2)^2 < 1$ by the AM-GM inequality. Thus $\tilde{Z}$ is obtained from Z' by adding a diagonal matrix with nonnegative entries and thus the matrix $\left[\begin{smallmatrix} Z_{00} & \mathbf{1}^T \\ \mathbf{1} & \tilde{Z} \end{smallmatrix}\right]$ is positive semidefinite.

Slater condition holds so the optimal value of (3) is equal to $\vartheta(G)$: indeed the choice $X = \frac{1}{2n} I_n$ and $x_i = 1/(2n)$ for all i is strictly feasible for (2), namely it satisfies the linear equality constraints and we have $\left[\begin{smallmatrix} 1 & x^T \\ x & X \end{smallmatrix}\right] \succ 0$. One way to prove the latter is to use the Schur complement lemma, stated below, and which we leave to the reader. □

Lemma 2 (Schur complement). *$\left[\begin{smallmatrix} A & B^T \\ B & C \end{smallmatrix}\right] \succ 0$ iff $A \succ 0$ and $C - BA^{-1}B^T \succ 0$.*

We now state a second lemma.

Lemma 3. *For any graph G we have $\vartheta(G)\vartheta(\bar{G}) \geq n$.*

Proof. Let Z be an optimal solution for (3) for the graph $G = (V, E)$. Let $X = Z/Z_{00} = Z/\vartheta(G)$ and note that X is feasible for (2) applied to $\bar{G}$. This shows that $\vartheta(\bar{G}) \geq \sum_i X_{ii} = n/\vartheta(G)$. □

We are now ready to prove Theorem 9.2

Proof of Theorem 9.2. Consider a random undirected graph G on $V = \{1, \ldots, n\}$ defined as follows: for each pair $\{i, j\} \subset V$ with $i \neq j$, we put an edge between i and j with probability $1/2$ and independently of the other pairs[1]. One can show that when n is large enough then $\alpha(G)$ is concentrated around $2\log(n)$, so that $\Pr[\alpha(G) \leq 3\log(n)] \geq 3/4$ for large enough n.

Note that G and $\bar{G}$ have the same distribution, so $\Pr[\alpha(\bar{G}) \leq 3\log(n)] \geq 3/4$. By the inclusion-exclusion principle we have

$$\Pr[\alpha(\bar{G}) \leq 3\log(n) \text{ and } \alpha(G) \leq 3\log(n)] \geq 3/4 + 3/4 - 1 > 0.$$

[1]Alternatively you can think of such a graph in terms of its adjacency matrix A, where each entry A_{ij} for $i < j$ is 0 with probability $1/2$ and 1 with probability $1/2$. Recall that the adjacency matrix of a graph $G = (V, E)$ is a matrix $A \in \mathbb{R}^{|V| \times |V|}$ where $A_{ij} = 1$ if i and j are connected by an edge and 0 otherwise.

This means that there exists at least one graph G such that $\max(\alpha(G), \alpha(\bar{G})) \leq 3\log(n)$. Since $\vartheta(G)\vartheta(\bar{G}) \geq n$ we get that one of $\vartheta(G) \geq \sqrt{n}$ or $\vartheta(\bar{G}) \geq \sqrt{n}$ is true. Thus inequality $\vartheta(G_n)/\alpha(G_n) \geq \sqrt{n}/(3\log n)$ is true for either $G_n = G$ or $G_n = \bar{G}$. $\square$

References

[BTN01] Aharon Ben-Tal and Arkadi Nemirovski. *Lectures on modern convex optimization: analysis, algorithms, and engineering applications.* SIAM, 2001. 2

[GW95] Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42(6):1115–1145, 1995.

[Lov79] László Lovász. On the Shannon capacity of a graph. *IEEE Transactions on Information Theory*, 25(1):1–7, 1979. 1