Stable Sets and Lovász's Theta Function
Fall 2012
Term Project

Instructor: Farid Alizadeh
Scribe: Emre Yamangil

12/10/2012

1 Overview

We will go over some of the classical valid inequalities for the stable set polyhedron and then provide the proof of Lovász's Theta function and show SDP techniques can be used to optimize over TH(G) in polynomial time.

2 Preliminaries on Stable Sets

Let $G = (V, E)$ be an undirected graph without isolated vertices. A stable set is a subset of nodes $A \subseteq V$ such that $ij \notin E$ for any $i, j \in A$. Let $\alpha(G)$ denote the size of maximum independent set within G. Let χ^A denote the characteristic vector of $A \subseteq V$.

$$\text{STAB}(G) = \text{conv}\{\chi^A : A \subseteq V \text{ a stable set}\}$$

Also let $P(G) = \{x \in \mathbb{R}^n_+ : x_i + x_j \leq 1, \forall ij \in E\}$. Therefore $\text{STAB}(G) = P(G) \cap \mathbb{B}^n$. $\text{STAB}(G) = P(G)$ if and only if G is bipartite. For arbitrary graphs we can easily show $P(G)$ is half-integral (that is, it's vertices are composed of $\{0, \frac{1}{2}, 1\}$).

3 Valid Inequalities for STAB(G)

3.1 Clique inequalities

An clique is a subset of nodes $B \subseteq V$ such that G_B (the graph induced by B) is a complete graph. Let $A \subseteq V$ be a stable set, obviously $|B \cap A| \leq 1$ (Otherwise there exists an edge within the stable set) or equivalently $\chi^B \bullet \chi^A \leq 1$, for any stable set A in G. Therefore we can conclude with the following set of valid inequalities for STAB(G):

$$\sum_{i \in B} x_i \leq 1, \quad \text{for any clique } B \subseteq V \tag{1}$$

Let

$$\text{QSTAB}(G) = P(G) \cap \{x : \sum_{i \in B} x_i \leq 1, \text{ for any clique } B \subseteq V\} \tag{2}$$

Definition 1 *A graph* G *is called perfect if* $\omega(G') = \alpha(G')$ *for all induced subgraphs* G' *of* G*.*

Bipartite graphs, line graphs of bipartite graphs and comparability graphs are perfect graphs. Another important consequence of perfect graphs is that the complement of a perfect graph is again a perfect graph (conjectured by Berge at 1961 and proved by Lovász by 1972). The following is an important characterization of perfect graphs:

Theorem 2 *STAB*$(G) =$ *QSTAB*(G) *if and only if* G *is perfect.*

3.2 Odd-hole inequalities

We call $C \subseteq V$ an odd-hole if it is a chordless odd cycle in G. Then the following inequality is obviously valid for STAB(G):

$$\sum_{i \in C} x_i \leq \frac{1}{2}(|C| - 1), \quad \text{for any odd hole } C \subseteq V \tag{3}$$

This can be generalized for odd cycles with chords, however they are essentially summation of edge inequalities ($x_i + x_j \leq 1$) and odd-hole inequalities.

Definition 3 *We call a graph* G, $t-$*perfect if STAB*$(G) = P(G) \cap \{x : \sum_{i \in C} x_i \leq \frac{1}{2}(|C| - 1),$ *for any odd hole* $C \subseteq V\}$*.*

Definition 4 *We call a graph* G, $h-$*perfect if STAB*$(G) = P(G) \cap \{x : \sum_{i \in C} x_i \leq \frac{1}{2}(|C| - 1),$ *for any odd hole* $C \subseteq V, \sum_{i \in B} x_i \leq 1,$ *for any clique* $B \subseteq V\}$*.*

3.3 Odd-antihole inequalities

We call $D \subseteq V$ an odd-antihole if it's complement is a chordless odd cycle in $\bar{G}$. Then the following inequality is obviously valid for STAB(G):

$$\sum_{i \in D} x_i \leq 2, \quad \text{for any odd antihole } D \subseteq V \tag{4}$$

Corollary 5 *If STAB*$(G) = P(G) \cap \{x : x \in (1) - (4)\}$*, then we can optimize over STAB*(G) *in polynomial time.*

This corollary follows from the fact that all these inequalities are rank 1 inequalities of the Lovász-Schrijver hierarchy ([2]).

3.4 Orthogonality inequalities

Definition 6 *An orthonormal representation of* G *is* $(u_i \in \mathbb{R}^N, i \in V)$ *such that* $\|u_i\| = 1$ *for all* i *and* $u_i \bullet u_j = 0$ *for* $ij \notin E$ *and* $N \in \mathbb{Z}_+$*. Obviously every graph has an orthonormal representation in* $\mathbb{R}^V$*.*

Let $S \subseteq V$ be a stable set in G and $c \in \mathbb{R}^N$ such that $\|c\| = 1$. Then:

$$\sum_{i \in S} (c^T u_i)^2 \leq 1$$

As vectors u_i, $i \in S$ are mutually orthogonal and by a rotation of the orthonormal representation we may assume they are unit vectors. Therefore $\sum_{i \in S} (c^T u_i)^2 \leq \|c\|^2 = 1$.

Since $\sum_{i \in V} (c^T u_i)^2 \chi_i^S = \sum_{i \in S} (c^T u_i)^2$ we may conclude:

$$\sum_{i \in V} (c^T u_i)^2 x_i \leq 1 \tag{5}$$

is valid for STAB(G).

Remark 7 *We note that (5) implies (1). To see this, let* $Q \subseteq V$ *be a clique in* G *and* $\{u_i : i \in V \setminus Q\} \cup \{c\}$ *be mutually orthogonal unit vectors. Set* $u_j = c$ *for* $j \in Q$*. This is obviously an orthonormal representation of* G*. Therefore we have* $1 \geq \sum_{j \in V} (c^T u_i)^2 x_i = \sum_{j \in Q} (c^T c)^2 x_i + \sum_{j \notin Q} (c^T u_i)^2 x_i = \sum_{j \in Q} x_i$.

4 Theta function

Let us start with the following polyhedron:

$$\text{TH}(G) = P(G) \cap \{x : \sum_{i \in V} (c^T u_i)^2 x_i \leq 1, (u_i, i \in V) \text{ an orth. rep. of } G, \|c\| = 1\}$$

Remark 7 implies that

$$\text{STAB}(G) \subseteq \text{TH}(G) \subseteq \text{QSTAB}(G)$$

Obviously TH(G) consists of infinitely many half-spaces, therefore it is convex. But it is not necessarily polyhedral. We will further characterize this towards the end of our discussion.

For a non-negative weight of vertices $w \in \mathbb{R}_+^n$ the theta function is defined as follows:

$$\theta(G,w) := \max\{w^T x : x \in \mathrm{TH}(G)\}$$

Let $\bar{w}_i = \sqrt{w_i}$, $i \in V$ and $W = \bar{w}\bar{w}^T$.

$$\begin{aligned}
\mathcal{F} &:= \{A \in \mathbb{R}^{n\times n} : A \text{ is symmetric}\}\\
\mathcal{M} &:= \{B \in \mathcal{F} : b_{ij} = 0, ij \in E\}\\
\mathcal{M}^\perp &:= \{A \in \mathcal{F} : A \bullet B = 0, B \in \mathcal{M}\}\\
\mathcal{D} &:= \{A \in \mathbb{R}^{n\times n} : A \text{ is positive semidefinite}\}
\end{aligned}$$

Let $\Delta(D)$ denote the largest eigenvalue of matrix D. We will define some other value functions which will be instrumental of the proof of main result.

$$\theta_1(G,w) := \min_{\{c,(u_i)\}} \max_{i\in V} \frac{w_i}{(c^T u_i)^2}$$

where $\|c\| = 1$ and $(u_i, i \in V)$ an orthonormal representation of G.

$$\theta_2(G,w) := \min\{\Delta(A+W) : A \in \mathcal{M}^\perp\}$$

$$\theta_3(G,w) := \max\{\bar{w}^T B \bar{w} : B \in \mathcal{D} \cap \mathcal{M}, \mathrm{tr}(B) = 1\}$$

$$\theta_4(G,w) := \max_{\{d,(v_i)\}} \sum_{i\in V} (d^T v_i)^2 w_i$$

where $\|d\| = 1$ and $(v_i, i \in V)$ an orthonormal representation of $\bar{G}$.

Theorem 8 $\theta(G,w) = \theta_1(G,w) = \theta_2(G,w) = \theta_3(G,w) = \theta_4(G,w)$.

Proof: Claim 1. $\theta(G,w) \le \theta_1(G,w)$. Let $x = \mathrm{argmax}\{w^T x : x \in \mathrm{TH}(G)\}$, $(u_i, i \in V)$ orthonormal representation of G and $\|c\| = 1$. Then

$$\begin{aligned}
\theta(G,w) = w^T x &= \sum_{i\in V} w_i x_i\\
&= \sum_{i\in V} w_i \frac{(c^T u_i)^2}{(c^T u_i)^2} x_i\\
&\le (\max_i \frac{w_i}{(c^T u_i)^2}) \sum_{i\in V} (c^T u_i)^2 x_i\\
&\le (\max_i \frac{w_i}{(c^T u_i)^2})
\end{aligned}$$

proves the claim (last inequality follows because $x \in \mathrm{TH}(G)$).

Claim 2. $\theta_1(G,w) \le \theta_2(G,w)$. Choose $A \in \mathcal{M}^\perp$ and set $t := \Delta(A+W) > 0$ (since $\sum_i \lambda_i(A+W) = \mathrm{tr}(A+W) = \mathrm{tr}(W) = \sum_{i\in V} w_i > 0$). Then $tI -$

$(A+W) \succcurlyeq 0$, implies $tI - (A+W) = X^TX$ for some matrix $X \in \mathbb{R}^{n\times n}$. Let x_i denote the ith column of X. Therefore $x_i^Tx_i = t - a_{ii} - w_i = t - w_i$ (as $A \in \mathcal{M}^\perp$), and $x_i^Tx_j = -\sqrt{w_iw_j}$ for $ij \notin E$. Let $||c|| = 1$ be orthogonal to all x_i, $i \in V$ (X is singular as one eigenvalue of $tI - (A+W)$ is equal to 0) and consider vectors $u_i := \sqrt{w_i/t}c + \sqrt{1/t}x_i$. Then we have $u_i^Tu_i = 1$ and $u_i^Tu_j = 0$. Therefore $(u_i, i \in V)$ is an orthonormal representation of G. As $c^Tu_i = \sqrt{w_i/t}c^Tc + \sqrt{1/t}c^Tx_i = \sqrt{w_i/t}$ we have

$$\theta_1(G,w) \le \max_{i\in V}\frac{w_i}{(c^Tu_i)^2} = \max_{i\in V}\frac{w_i}{\sqrt{w_i/t}^2} = t = \Delta(A+W)$$

proves the claim.

Claim 3. $\theta_2(G,w) \le \theta_3(G,w)$. Let $\theta_3 := \theta_3(G,w)$ and B feasible with respect to $\theta_3(G,w)$. Then we have $\bar{w}^TB\bar{w} \le \theta_3\operatorname{tr}(B)$ or equivalently

$$\begin{aligned}\bar{w}^TB\bar{w} - \theta_3\operatorname{tr}(B) &= (\bar{w}\bar{w}^T)\bullet B - \theta_3 I \bullet B \\ &= (\bar{w}\bar{w}^T - \theta_3 I)\bullet B \le 0 \end{aligned} \tag{6}$$

As $\mathcal{D}$ and $\mathcal{M}$ are cones $\mathcal{D}\cap\mathcal{M}$ is also a cone. Let us define the polar of a cone as follows, $K^\circ = \{x : x \bullet y \le 0, y \in K\}$. Therefore by (6):

$$W - \theta_3 I \in (\mathcal{D}\cap\mathcal{M})^\circ = \mathcal{D}^\circ + \mathcal{M}^\circ = -\mathcal{D} + \mathcal{M}^\perp$$

Let $D \in \mathcal{D}$ and $-A \in \mathcal{M}^\perp$ such that $W - \theta_3 I = -D - A$ or $D = \theta_3 I - (A+W) \in \mathcal{D}$. Therefore $\theta_3 I \succcurlyeq A + W$ implying $\theta_3 \ge \Delta(A+W) \ge \theta_2(G,w)$, proves the claim.

Claim 4. $\theta_3(G,w) \le \theta_4(G,w)$. Let B optimal with respect to $\theta_3(G,w)$ ($\bar{w}^TB\bar{w} = \theta_3$). As B is positive semidefinite by feasibility we have $B = Y^TY$ for some matrix $Y \in \mathbb{R}^{n\times n}$. Let y_i denote the ith column of Y. Let $P := \{i \in V : y_i \neq 0\}$. Set $v_i := \frac{1}{||y_i||}y_i$ for $i \in P$. For $i \in V\setminus P$ choose an orthonormal basis of the linear space $(\operatorname{lin}\{v_i : i \in P\})^\perp$ and use its elements for v_i. As $v_i^Tv_j = \frac{y_i^Ty_j}{||y_i||||y_j||} = \frac{b_{ij}}{||y_i||||y_j||} = 0$ if $ij \in E$, $(v_i, i \in V)$ forms an orthonormal representation of $\bar{G}$. Let $d := \frac{Y\bar{w}}{\sqrt{\theta_3}}$ is a unit length vector. Let us consider

$$d^Tv_i = \frac{\bar{w}^TY^TYe_i}{\sqrt{\theta_3}||y_i||} = \frac{\bar{w}Be_i}{\sqrt{\theta_3}||y_i||}$$

In other words $||y_i||d^Tv_i = \frac{\bar{w}^TBe_i}{\sqrt{\theta_3}}$. Therefore

$$\sum_{i\in V}||y_i||d^Tv_i\sqrt{w_i} = \frac{1}{\sqrt{\theta_3}}\sum_{i\in V}\bar{w}^TBe_i\sqrt{w_i} = \frac{\bar{w}^TB\bar{w}}{\sqrt{\theta_3}} = \sqrt{\theta_3}$$

All together we have the following implication:

$$\begin{aligned}\theta_3 &= (\sum_{i\in V} ||y_i|| d^T v_i \sqrt{w_i})^2 \\ &\leq (\sum_{i\in V} ||y_i||^2)(\sum_{i\in V} (d^T v_i)^2 w_i) \\ &= \mathrm{tr}(B)(\sum_{i\in V} (d^T v_i)^2 w_i) \\ &= \sum_{i\in V} (d^T v_i)^2 w_i \leq \theta_4(G, w)\end{aligned}$$

first step involves Cauchy-Schwarz inequality and we also use $B_{ii} = y_i^T y_i = ||y_i||^2$.

Claim 5. $\theta_4(G,w) \leq \theta(G,w)$. Choose $(v_i, i \in V)$ an orthonormal representation of $\bar{G}$ and $||d|| = 1$ such that $\theta_4(G,w)$ is achieved. Let $(u_i, i \in V)$ be an orthonormal representation of G with $||c|| = 1$. Let us consider the matrices $u_i v_i^T \in \mathbb{R}^{n\times n}$. As $u_i v_i^T \bullet u_i v_i^T = (u_i^T u_i)(v_i^T v_i) = 1$, $u_i v_i^T$ is of unit length. Similarly the matrix cd^T is also of unit length. Moreover, $u_i v_i^T \bullet u_j v_j^T = (u_i^T u_j)(v_i^T v_j) = 0$ as either $u_i^T u_j = 0$ or $v_i^T v_j = 0$ (the edge cannot be present both in G and $\bar{G}$). Therefore $(U_i = u_i v_i^T, i \in V)$ are mutually orthogonal with $C = cd^T$ s.t. $||C|| = 1$. This implies

$$\sum_{i\in V} (C \bullet U_i)^2 \leq 1$$

with the same reasoning used in the orthogonal valid inequalities. Therefore

$$1 \geq \sum_{i\in V} (cd^T \bullet u_i v_i^T)^2 = \sum_{i\in V} (c^T u_i)^2 (d^T v_i)^2$$

implies $((d^T v_i)^2, i \in V) \in \mathrm{TH}(G)$. $\theta_4(G,w) = \sum_{i\in V} w_i (d^T v_i)^2 \leq \theta(G,w)$ as a result of feasibility, completes the proof of theorem. ■

Corollary 9 *A linear function can be optimized over TH*(G) *in polynomial time.*

This is a direct corollary to the Theorem 8 as $\theta_2(G,w)$ is a well-known Semidefinite Program (SDP) which can be solved in polynomial time in the size of graph G.

Theorem 10 *All facets of TH*(G) *are facets of QSTAB*(G)*.*

Proof: Let, $F = \{x \in \mathrm{TH}(G) : \sum_{i\in V} a_i x_i = \alpha\}$ be a facet of $\mathrm{TH}(G)$. As $0 \in \mathrm{TH}(G)$, we may assume $\alpha = 1$. Therefore $F = \{x \in \mathrm{TH}(G) : \sum_{i\in V} a_i x_i = 1\}$.

As F is a facet, there exists $z \in \text{int}(F)$. Let $(u_i, i \in V)$ be an orthonormal representation of G and a $\|c\| = 1$ s.t. $a_i = (c^T u_i)^2$. Therefore $\sum_{i \in V}(c^T u_i)^2 z_i = 1 = \theta(\bar{G}, z)$ (By Corollary 9.3.22). We have the following:

$$\sum_{i \in V}(c^T u_i)^2 z_i = \theta(\bar{G}, z)$$
$$\sum_{i \in V}(\bar{c}^T u_i)^2 z_i \leq \bar{c}^T \theta(\bar{G}, z)\bar{c}$$
$$\bar{c}^T(\sum_{i \in V} z_i u_i u_i^T)\bar{c} \leq \bar{c}^T \theta(\bar{G}, z)\bar{c}$$

The quantity on the left is maximized by an eigenvector of $(\sum_{i \in V} z_i u_i u_i^T)$, namely c, and it's maximum value is it's corresponding eigenvalue $\theta(\bar{G}, z)$. Therefore we have

$$(\sum_{i \in V} z_i u_i u_i^T)c = \theta(\bar{G}, z)c$$
$$\sum_{i \in V} z_i(c^T u_i)u_i = c$$
$$\sum_{i \in V} z_i(c^T u_i)(u_i)_j = c_j, \ j = 1, \ldots, N \tag{7}$$

As F is of $n-1$ dimensional (Assuming TH(G) is full dimensional), (7) should be all consequences of $\sum_{i \in V} a_i x_i = 1$. Therefore $(c^T u_i)u_i = (c^T u_i)^2 c$. If $c^T u_i \neq 0$ then $(u_i) = (c^T u_i)c$ implies $u_i = \pm c$ (we may assume $u_i = c$). Therefore either $c^T u_i = 0$ or $u_i = c$. If we let $Q = \{i : u_i = c\}$, then the clique inequality is equivalent with $\sum_{i \in V} a_i x_i = 1$. ∎

Corollary 11 *TH*(G) *is polyhedral if and only if* G *is perfect.*

Bibliography

References

[1] M. Grotschel, L. Lovasz, and A. Schrijver, Geometric Algorithms and Combinatorial Optimization, Springer-Verlag, Berlin, New York.

[2] L. Lovasz, and A. Schrijver, Cones of Matrices and Set-functions and 0-1 Optimization, SIAM Journal of Optimization, Vol 1(2) (1991) 166-190.