Semidefinite Programming	**Lecture 8**
OR 6327 Spring 2012	**February 16, 2012**
Scribe: Daniel Fleischman	

1 Shannon Capacity - Motivation

Suppose you have a channel where you can send one "symbol" per time interval, from a finite set of n symbols. Upon proper encoding, without any further constraint, you would be able to transmit $\log n$ bits of information per time interval (all logs are base 2).

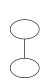
The problem is that some of the symbols are very much alike, and their difference may be lost during the communication. To formalize this, let's say that there is a graph $G = (V, E)$ where V is the set of nodes, and $ij \in E$ if symbols i and j can be confused.

One way to solve this problem is by selecting a stable set S (i.e., $ij \notin E$ for all $i, j \in S$) and only send symbols from this set, since they can't be confused with each other. If the two sides of the communication channel know what is the set S, we can send $\log |S|$ bits of information per time interval. The size of the maximum (cardinality) stable set in G is called $\alpha(G)$, and we have a limit of $\log \alpha(G)$ bits per time interval.

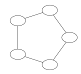
But we can certainly try to do something smarter. Suppose that instead of symbols, we communicate by "pairs of symbols". Two pairs of symbols can be confused if each of the two elements can be confused (or are equal). To make it formal, we will define the graph $G^2 = (V \times V, E^2)$ where $(ik, jl) \in E$ if ($i = j$ or $ij \in E$) and ($k = l$ or $kl \in E$). Now, if we find any stable set S of G^2 we can send an element of S per two time intervals, giving a limit of $\frac{1}{2} \log \alpha(G^2)$ bits per time interval.

Of course we can generalize this to get something called the **Shannon Capacity**. We define $\rho(G) = \sup_k \frac{1}{k} \log \alpha(G^k)$, and $\Theta(G) = 2^{\rho(G)} = \sup_k \sqrt[k]{\alpha(G^k)}$. $\Theta(G)$ is the Shannon Capacity of G.

2 Examples

Consider $G = K_2 =$  $\Rightarrow G^k = K_{2^k}$, and $\alpha(G^k) = 1$, $\rho(G) = 0$ and $\Theta(G) = 1$.

Now consider $G = \bar{K}_n$ the graph with n vertices and no edges. $G^k = \bar{K}_{n^k}$, $\alpha(G^k) = n^k$, $\rho(G) = \log n$ and $\Theta(G) = n$, as expected (if there is no confusion, we can send $\log n$ bits of information).

Finally, let $G = C_5 =$ . Clearly $\alpha(G) = 2$. Also, $\alpha(G^2) = 5$ (e.g. aa, bc, ce, ed, db), therefore $\rho(G) \geq \frac{1}{2}\log 5$ and $\Theta(G) \geq \sqrt{5}$. This is actually the value of $\Theta(C_5)$ (as we will see). As a note, $\Theta(C_7)$ is still unknown.

3 The Lovász Theta Function

In 1979 Lovász introduced a function of a graph $\vartheta(G)$ which can be computed in polynomial time. He also proved that $\Theta(G) \leq \vartheta(G)$. Also, $\vartheta(C_5) \leq \sqrt{5}$, giving $\Theta(C_5) = \vartheta(C_5) = \sqrt{5}$. To define $\vartheta(G)$ we need another definition first.

Definition 1 $(c, u_1, \ldots, u_n)$ *is an* **orthonormal representation** *of a graph* $G = (N, E)$, $N = \{1, \ldots, n\}$ *if* $c, u_1, \ldots, u_n \in \mathbb{R}^n$ *are unit vectors with* $u_i^T u_j = 0$ *if* $ij \notin E$.

As an example, suppose $G = K_n$. Then any set of unit vectors, e.g., $c = u_1 = \ldots = u_n$ is an orthonormal representation. On the other hand, if $G = \bar{K}_n$ then $u_1, \ldots, u_n$ should form an orthonormal basis of $\mathbb{R}^n$. For instance, $u_i = e_i$ and $c = \frac{e}{\sqrt{n}}$.

Definition 2 *The* **Lovász Theta Function** *of a graph is defined as:*

$$\vartheta(G) = \min_{(c,u_1,\ldots,u_n)} \max_{1 \leq i \leq n} \frac{1}{(c^T u_i)^2}$$

where the minimum ranges over all orthonormal representations of G.

Note that $\vartheta(G) \geq 1$, since $|c^T u_i| \leq 1$. Also, note that the example of orthonormal representation that we gave for $\bar{K}_n$ is valid for all possible graphs, and has a value of n. Therefore, $1 \leq \vartheta(G) \leq n$.

Remark 1 *Since the optimal orthonormal representation has* $|c^T u_i| \geq 1/\sqrt{n}$ *for all* i*, we can restrict ourselves to these. Then we are minimizing a continuous function over a compact set, therefore the minimum is attained.*

4 Main Result

The goal of the rest of this lecture is to prove that $\Theta(G) \leq \vartheta(G)$. We will prove this via three lemmas.

Definition 3 *Suppose that* $G = (V, E)$ *and* $H = (W, F)$ *are graphs. Then* $G \times H$*, the* **strong product** *of* G *and* H *is defined on the node set* $V \times W$ *by two nodes* $v_1 w_1$ *and* $v_2 w_2$ *being adjacent if* $(v_1 = v_2$ *or* $v_1 v_2 \in E)$ *and* $(w_1 = w_2$ *or* $w_1 w_2 \in F)$.

Lemma 1 *$\alpha(G \times H) \geq \alpha(G)\alpha(H)$.*

Proof: If S and T are maximum stable sets of G and H, then $S \times T$ is a stable set of $G \times H$. $\square$

Corollary 1 *$\alpha(G^k) \geq (\alpha(G))^k$, so $\Theta(G) = \sup_k \sqrt[k]{\alpha(G^k)} = \lim_k \sqrt[k]{\alpha(G^k)}$.*
$\square$

Lemma 2 *$\vartheta(G \times H) \leq \vartheta(G)\vartheta(H)$.*

Proof: Let $(c, u_1, \ldots, u_m)$ and $(d, v_1, \ldots, v_n)$ be orthonormal representations of G and H attaining $\vartheta(G)$ and $\vartheta(H)$. To find a representation of $G \times H$ we use the Kronecker product, which when applied to vectors can be defined as $u \otimes v = \text{vec}(uv^T)$. Note that:

$$(u \otimes v)^T(\bar{u} \otimes \bar{v}) = \text{vec}(uv^T)^T \text{vec}(\bar{u}\bar{v}^T) = uv^T \bullet \bar{u}\bar{v}^T = \text{trace}\,(vu^T\bar{u}\bar{v}^T) = (u^T\bar{u})(v^T\bar{v}).$$

Hence, $(c \otimes d, u_i \otimes v_j, 1 \leq i \leq m, 1 \leq j \leq n)$ is an orthonormal representation of $G \times H$. So:

$$\begin{aligned}
\vartheta(G \times H) &\leq \max_{i,j} \frac{1}{[(c \otimes d)^T(u_i \otimes v_j)]^2} \\
&= \max_{i,j} \frac{1}{(c^Tu_i)^2(d^Tv_j)^2} \\
&= \vartheta(G)\vartheta(H).
\end{aligned}$$

$\square$

Lemma 3 *$\vartheta(G) \geq \alpha(G)$.*

Proof: Let S be the maximum stable set of G, say (w.l.o.g.) $S = \{1, \ldots, k\}$, with $k = \alpha(G)$. Let $(c, u_1, \ldots, u_n)$ be any orthonormal representation of G. Then $u_1, \ldots, u_k$ are orthonormal. Extend it with vectors $v_{k+1}, \ldots, v_n$ such that $u_1, \ldots u_k, v_{k+1}, \ldots, v_n$ is an orthonormal basis of $\mathbb{R}^n$. Then:

$$\begin{aligned}
1 = \|c\|^2 &= \sum_{i=1}^{k}(c^Tu_i)^2 + \sum_{j=k+1}^{n}(c^Tv_j)^2 \\
&\geq \sum_{i=1}^{k}(c^Tu_i)^2 \\
&\geq k \min_i (c^Tu_i)^2.
\end{aligned}$$

Therefore, $\max_i \frac{1}{(c^Tu_i)^2} \geq k$ and $\vartheta(G) \geq \alpha(G)$. $\square$

Finally, we have the theorem.

Theorem 1 *$\Theta(G) \leq \vartheta(G)$.*

Proof: For any k we have that $\alpha(G^k) \leq \vartheta(G^k) \leq \vartheta(G)^k$. So $\sqrt[k]{\alpha(G^k)} \leq \vartheta(G)$ and $\Theta(G) \leq \vartheta(G)$. $\square$