PRINCETON UNIVERSITY
ORFE

Instructor:
Amir Ali Ahmadi
Spring 2016

TA: Georgina Hall

This lecture:

Some applications of SDP in combinatorial optimization:

1. upper bounding the stability number of a graph
   - SDP upper bound of Lovász
   - LP upper bounds with clique inequalities
2. Upperbounding the Shannon capacity of a graph.

Fort part 1 of the lecture, please refer to Georgina's scribed notes.

I only outline it here:

$\alpha(G) :=$ size of the largest independent (aka stable) set in a graph $G(V,E)$ on $n$ nodes.

$$\vartheta(G) := \max_{X \in S^{n \times n}} \operatorname{Tr}(JX) \quad \text{s.t.} \quad \operatorname{Tr}(X) = 1, \quad X_{ij} = 0 \text{ if } (i,j) \in E, \quad X \succeq 0$$

$J$: matrix of all ones.

The Lovasz SDP [Lov79]

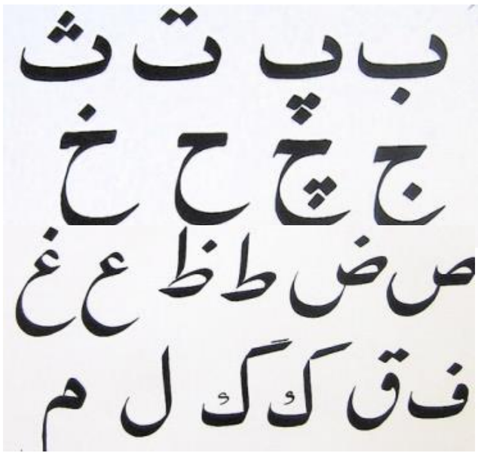
I proved in class that $\alpha(G) \leq \vartheta(G)$. In your homework you also prove that $\vartheta(G) \leq \chi(\bar{G})$, where $\bar{G}$ is the complement graph of $G$ and $\chi$ denotes the coloring number

We also saw a family of LP-based upperbounds on $\alpha(G)$, which on your homework you will prove are all dominated by the SDP bound $\vartheta(G)$.

## Upper bounds on the Shannon capacity of a graph

The Lovasz SDP that we presented previously was in fact introduced to tackle a (rather difficult) problem in coding theory, put forward by Claude Shannon [Sha56].

Suppose you have an alphabet with a finite number of letters $v_1, \dots, v_m$. You want transmit messages from this alphabet over a noisy channel. Some of your letters look similar and can get confused at the receiver end because of noise.




Some letters from the Persian alphabet.

Think for example of the letters ک and گ depicted here.

Consider a graph $G$ whose nodes are the letters and which has an edge between two nodes iff the two letters can get confused. How many 1-letter words can we send from our alphabet so that we are guaranteed to have no confusion at the receiver? Well, this would be exactly $\alpha(G)$, the stability number of the graph.

But how many 2-letter words can we send with no confusion? How many 3-letter words? And so on...

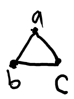
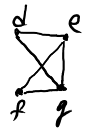
Note that two $k$-letter words can be confused iff each of their letters can be confused or are equal. Hence, the number of $k$-letter words that can be sent without confusion is exactly

$$\alpha(G^k),$$

where $G^k := \underbrace{G \otimes G \otimes \cdots \otimes G}_{k \text{ times}}$,

and $\otimes$ denotes the strong graph product defined as follows

**Definition.** Consider two graphs $G_A(V_A, E_A)$ and $G_B(V_B, E_B)$, with $|V_A| = n$ and $|V_B| = m$. Then their strong graph product $G_A \otimes G_B$ is a graph with $nm$ nodes $V_A \times V_B$, where two nodes $(i,k)$ and $(j,l)$ are connected iff $(i - j$ is an edge in $G_A$ or $i = j)$ and $(k - l$ is an edge in $G_B$ or $k = l)$.

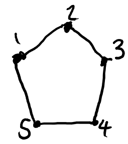
**Practice:** Draw $G_A \otimes G_B$ if $G_A =$  and $G_B =$ .

**Lemma 2.1.** $\alpha(G_A \otimes G_B) \geq \alpha(G_A) \cdot \alpha(G_B)$. (In particular, $\alpha(G^k) \geq \alpha^k(G)$.)

**Proof.** Let $S_1 = \{u_1, \ldots, u_l\}$ be a maximum stable set in $G_A$ and $S_2 = \{v_1, \ldots, v_r\}$ be a maximum stable set in $G_B$. Then,
$S_1 \times S_2 := \{u_i v_j : u_i \in S_1, v_j \in S_2\}$ is a stable set in $G_A \otimes G_B$.

It is quite possible, however, to have $\alpha(G_A \otimes G_B) > \alpha(G_A) \cdot \alpha(G_B)$.

Here is an example:

Let $G_A = G_B = C_5$ (i.e., a cycle on five nodes  ).



Then $\alpha(G_A) \cdot \alpha(G_B) = 2 \cdot 2 = 4$, but $\alpha(G_A \otimes G_B) = 5$ since

the set $\{a_1b_1, a_2b_3, a_3b_5, a_5b_4, a_4b_2\}$ e.g. is a stable set in $G_A \otimes G_B$.

Definition. The Shannon capacity of a graph $G$, denoted by $\Theta(G)$, is defined as

$$\Theta(G) = \lim_{k \to \infty} \alpha^{1/k}(G^k).$$

One can show (e.g., by using Fekete's lemma) that the limit always exists and be equivalently written as

$$\Theta(G) = \sup_k \alpha^{1/k}(G^k).$$

Fekete's lemma. Consider a sequence $\{a_k\}$ that is superadditive; i.e., satisfies

$$a_{n+m} \geq a_n + a_m \quad \forall m,n.$$

Then, $\lim_{k \to \infty} \frac{a_k}{k}$ exists (in $\mathbb{R} \cup \{+\infty\}$) and is equal to $\sup_k \frac{a_k}{k}$.

To see the claim for $\Theta(G)$, apply this lemma to the sequence $\{\log \alpha(G^k)\}$ and observe that the strong graph product is associative.

The definition of $\Theta(G)$ itself suggest a natural way of obtaining lower bounds: for any $k$, we have

$$\alpha^{1/k}(G^k) \leq \Theta(G).$$

But how can we obtain an upper bound?

Note that for a graph $G$ with $n$ vertices, we always have

$$\Theta(G) \leq n,$$

because clearly $\alpha(G^k) \leq n^k \ \forall k$. But an upper bound of $n$ is almost always very loose.

In 1979, Lovasz [Lov79] gave an algorithm for computing upper bounds on the Shannon capacity that resolved the exact value for many more graphs than known before. In the process, he invented semidefinite programming. (Of course, he didn't call it that.)

We now give a proof of his result — that $\Theta(G) \leq \vartheta(G)$ — in the language of SDP.

<u>Theorem</u> 2.2. [Lov79] $\Theta(G) \leq \vartheta(G)$.

Here $\vartheta(G)$ is the optimal solution of the SDP seen before

$$\vartheta(G) := \max_X \ Tr\, JX \qquad \text{SDP 1}$$

$$\text{s.t. } Tr(X)=1, \ X \succeq 0$$

$$X_{ij}=0, \quad i,j \in E.$$

Recall that we have already shown that for any graph $G$

- $\alpha(G) \leq \vartheta(G)$ ①

- $\Theta(G) = \sup_k \alpha^{1/k}(G^k)$ ②

So if we also show that

$$\vartheta(G^k) \leq \vartheta^k(G), \quad ③$$

then we get

$$\Theta(G) \overset{②}{=} \sup_k \alpha^{1/k}(G^k) \overset{①}{\leq} \sup_k \vartheta^{1/k}(G^k) \overset{③}{\leq} \sup_k \vartheta(G) = \vartheta(G),$$

hence proving Theorem 2.2. The inequality in ③ is a direct corollary of the following theorem.

<u>Theorem 2.3</u>. For any two graphs $G_A$ and $G_B$ we have

$$\vartheta(G_A \otimes G_B) \leq \vartheta(G_A) \cdot \vartheta(G_B).$$

To prove this theorem, we need to take a feasible solution to <u>SDP1</u> applied to $G_A \otimes G_B$ and from it feasible solutions to <u>SDP1</u> applied to $G_A$ and $G_B$. This doesn't seem like a straightforward thing to do since we should somehow apply a "reverse Kronecker product" operation to our original solution. It would have been much nicer if we could turn the feasibility implication around in the other direction. Indeed we can, by taking the dual!

$$\max \quad \mathrm{Tr}\, JX \qquad \underline{\text{SDP1}}$$
$$\mathrm{Tr}(X) = 1$$
$$X_{ij} = 0, \quad i,j \in E$$
$$X \succeq 0$$

What is the dual of this SDP? Recall the standard primal-dual pair that we derived before:

(P) $\min_{X \in S^{n \times n}} \mathrm{Tr}(CX)$ $\quad$ (D) $\max_{y \in \mathbb{R}^m} b^T y$

$\mathrm{Tr}(A_i X) = b_i \quad i = 1, \dots, m$ $\quad$ $\sum_{i=1}^m y_i A_i \preceq C.$

$X \succeq 0$

Just by pattern matching, we see the dual of $\underline{\text{SDP1}}$ to be:

$$-\max_{t, y_{ij}} t \quad \xleftrightarrow{\text{dual}} \quad -\min \mathrm{Tr}(-JX)$$

$$tI + \sum y_{ij} E_{ij} \preceq -J \qquad \mathrm{Tr}(X) = 1, X_{ij} = 0, i,j \in E, X \succeq 0$$

$$\updownarrow$$

$$\min_{t, y_{ij}} -t \quad \longleftrightarrow \quad \min_{t, y_{ij}} t \qquad \underline{\text{SDP1d}}$$

$$-tI - \sum y_{ij} E_{ij} \succeq J \qquad tI + \sum y_{ij} E_{ij} \succeq J,$$

Where $E_{ij}$ with a matrix with a one in $(i,j)$ and $(j,i)$ position and zero o.w.

Let us rewrite sdp2d slightly:

$$\begin{aligned} \max_{t,A} \quad & t \\ \text{s.t.} \quad & tI + A - J \preceq 0 \\ & A_{ij} = 0, \text{ if } i=j \text{ or } (i,j) \notin E. \end{aligned}$$

Unlike LPs, SDPs do not always enjoy the property of having zero duality gap. However, since SDP1 and SDP2d are both strictly feasible there is indeed no duality gap. Hence:

$$\boxed{\begin{aligned} \vartheta(G) = \min_{t \in \mathbb{R}, Z \in S^{n \times n}} \quad & t \\ \text{s.t.} \quad & tI + Z - J \succeq 0 \\ & Z_{ij} = 0, \text{ when } i=j \text{ or } (i,j) \notin E. \end{aligned}} \qquad \text{(SDP2d)}$$

We now show that $\vartheta(G_A \otimes G_B) \leq \vartheta(G_A)\vartheta(G_B)$. First we prove a simple lemma from linear algebra.

Lemma 2.4. Consider two matrices $X \in S^{n \times n}$, $Y \in S^{m \times m}$.

$$X \succeq 0 \text{ and } Y \succeq 0 \Rightarrow X \otimes Y \succeq 0,$$

where $\otimes$ denotes the matrix Kronecker product.

Proof. Let $X$ have eigenvalues $\lambda_1, \ldots, \lambda_n$ with eigenvectors $v_1, \ldots, v_n$ and

$Y$ have eigenvalues $\mu_1, \dots, \mu_m$ with eigenvalues $w_1, \dots, w_m$. Then

$$(X \otimes Y)(v_i \otimes w_j) \overset{\text{verify!}}{=} X v_i \otimes Y w_j$$

$$= \lambda_i \mu_j \, v_i \otimes w_j.$$

Hence $X \otimes Y$ has $mn$ eigenvalues given by $\lambda_i \mu_j$, $i = 1, \dots, n$, $j = 1, \dots, m$, which are all nonnegative. $\square$

<u>Proof of Thm. 2.3</u>: Let $G_A$ and $G_B$ graphs on $n$ and $m$ nodes respectively. Consider a feasible solution $(t_A \in \mathbb{R}, A \in S^{n \times n})$ to SDP2d for $G_A$ and $(t_B \in \mathbb{R}, B \in S^{m \times m})$ to SDP2d for $G_B$.

We claim that the pair $(t_A t_B, C)$ with

$$C := t_A I_n \otimes B + t_B A \otimes I_m + A \otimes B$$

is feasible for SDP2d applied to $G_A \otimes G_B$ (and obviously has objective value $t_A t_B$). This would finish the proof. By assumption we have

$$t_A I_n + A - J_n \succeq 0 \Rightarrow t_A I_n + A + J_n \succeq 0 \quad (\text{b/c } 2J_n \text{ is psd.})$$

$$t_B I_m + B - J_m \succeq 0 \Rightarrow t_B I_m + B + J_m \succeq 0.$$

By Lemma 2.4 we have

$$(t_A I_n + A - J_n) \otimes (t_B I_m + B + J_m) \succeq 0 \qquad ①$$

$$(t_A I_n + A + J_n) \otimes (t_B I_m + B - J_m) \succeq 0 \qquad ②$$

$$① \Rightarrow (t_A I_n + A) \otimes (t_B I_m + B) + \cancel{(t_A I_n + A) \otimes J_m} - \cancel{J_n \otimes (t_B I_m + B)} - J_n \otimes J_m \succeq 0$$

$$② \Rightarrow (t_A I_n + A) \otimes (t_B I_m + B) - \cancel{(t_A I_n + A) \otimes J_m} + \cancel{J_n \otimes (t_B I_m + B)} - J_n \otimes J_m \succeq 0$$

Averaging both LMIs, we get

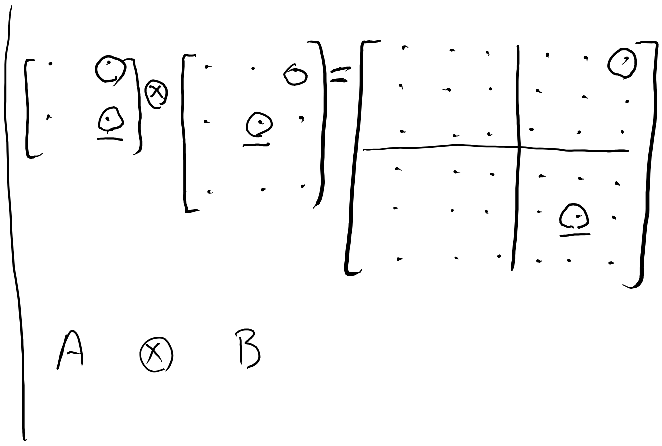
$$(t_A I_n + A) \otimes (t_B I_m + B) - J_n \otimes J_m \succeq 0.$$

$$\Rightarrow \underbrace{t_A t_B I_{nm} + t_A I_n \otimes B + t_B A \otimes I_m + A \otimes B}_{C} - J_n J_m \succeq 0.$$

Hence the required LMI is met. Lastly, we need to check that $C_{ij} = 0$ if $ij \notin E$, or if $i = j$.

Let us reindex $(i,j)$ in $G_A \otimes G_B$ as $(\tilde{i},\tilde{j}),(k,l)$, where $\tilde{i},\tilde{j}$ are nodes in $G_A$ and $k,l$ are nodes in $G_B$. The fact that there is no edge between the super node $(\tilde{i},k)$ and $(\tilde{j},l)$ in $G_A \otimes G_B$ means that either $\tilde{i}-\tilde{j}$ is not an edge in $G_A$ or $k-l$ is not an edge in $G_B$ or both.

- First observe that the $C_{ii} = 0$ because $A_{ii} = 0$ and $B_{ii} = 0$.



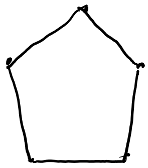

- Now consider $(i,j) = (\tilde{i},\tilde{j}),(k,l) \notin E$.

$$C_{(\tilde{i},\tilde{j}),(k,l)} = A_{\tilde{i}\tilde{j}} B_{kl} + t_A I_{\tilde{i}\tilde{j}} B_{kl} + t_B A_{\tilde{i}\tilde{j}} I_{k,l}$$

Either $A_{\tilde{i},\tilde{j}}$ or $B_{k,l}$ must be zero. Let's say w.l.o.g $B_{kl} = 0$. So we only have to worry about $A_{\tilde{i}\tilde{j}} I_{k,l}$.

If $k \neq l$, then $I_{k,l} = 0$ and we are done. If $k = l$, then $\tilde{i}-\tilde{j}$ cannot be an edge or else the letters would get confused (and hence $(\tilde{i},\tilde{j}),(k,l) \in E$). Hence, if $k = l$, we must have $A_{\tilde{i},\tilde{j}} = 0$. $\square$

**Example 1.** $\Theta(C_5) = ?$

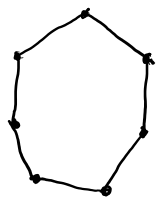


- $\alpha^{1/2}(C_5^2) = \sqrt{5} \leq \Theta(C_5)$.
- $\Theta(C_5) \leq \vartheta(C_5) = \sqrt{5}$.

$\Rightarrow \Theta(C_5) = \sqrt{5} = 2.2361\ldots$

Lovasz [Lov79] settled the exact value of $\Theta(C_5)$ more than 20 years after Shannon's paper [Sha59].

**Example 2.** $\Theta(C_7) = ?$



This is an open problem!

We know

$$\alpha^{1/5}(C_7^5) \leq 3.2271 \leq \Theta(C_7) \leq 3.3177 = \vartheta(C_7).$$

- The exact value of $C_7$ = An automatic A in ORF 523.

- Showing that $\Theta(C_7) < \vartheta(C_7)$ (if true) = Automatic 100/100 on the final exam.

Notes

Further reading for thi lecture can include chapter 2 of [LV12] and Chapter 3 of [BN01].

References

[Sha56] C.E. Shannon. The zero-error capacity of a noisy channel. IRE Trans. Inform. Theory, vol. IT2, no. 3, pp. 8-19, 1956.

[Lov79] L. Lovász. On Shannon capacity of a graph. IEEE Trans. on Inf. theory, 1979.

[LV12] M. Laurent and F. Vallentin. Lecture Notes on Semidefinite Optimization, 2012.

[BN01] A. Ben-Tal and A. Nemirovski. Lectures on Modern Convex Opt., SIAM, 2001.