Applications to Combinatorial Optimization

We are going to talk about two applications:

- The Lovász theta function (this lecture).
- The maxcut problem (next lecture).

(For more, see Goemans’s survey paper on SDP in Combinatorial Optimization.)

The Lovász \( \vartheta \) function

Let \( G = (V, E) \) be an undirected graph, with no loops or multiple edges. Let \(|V| = n, |E| = m\).

Definition 1 A stable set of \( G \) is a set of mutually nonadjacent vertices.

Definition 2 A clique of \( G \) is a set of mutually adjacent vertices.

Example 1 (pentagon)

\[
\begin{align*}
C_5 &  \\
1 & \\
5 \quad 2 & \\
4 & 3
\end{align*}
\]

\(\{1,3\}\) is a stable set;
\(\{1,2\}\) is a clique.

There are two quantities we are interested in:

\[
\begin{align*}
\alpha(G) & := \text{maximum cardinality of a stable set of } G. \\
\chi(G) & := \text{minimum cardinality of a clique cover of } G.
\end{align*}
\]

In the pentagon example, \( \alpha(G) = 2 \) (e.g. \( \{1,3\}\)), and \( \chi(G) = 3 \) (e.g. \( \{1,2\}, \{3,4\}, \{5\}\)).

These two quantities are closely related. In particular, we always have \( \alpha(G) \leq \chi(G) \). To see why this is true, consider any stable set. Then by definition of a stable set, the elements of
this set must all be in different sets in any clique cover.

Both these quantities are NP-hard to compute. But there is a graph invariant $\vartheta(G)$ such that

$$\alpha(G) \leq \vartheta(G) \leq \chi(G),$$

and $\vartheta(G)$ is the optimal value of an SDP problem (Lovász, '79).

We will now define $\vartheta$ and show the first inequality.

(Notation: $e = (1; 1; \ldots; 1)$, $0/0 := 0$)

Where does the SDP come from?

$$\alpha(G) = \max_x \left\{ e^T x : x_i x_j = 0 \forall ij \in E, \ x \in \{0,1\}^n \right\}$$

$$= \max_x \left\{ \frac{(e^T x)^2}{x^T x} : x_i x_j = 0 \forall ij \in E, \ x \in \{0,1\}^n \right\}$$

$$\leq \max_x \left\{ \frac{(e^T x)^2}{x^T x} : x_i x_j = 0 \forall ij \in E, \ x \geq 0 \right\}$$

$$= \max_x \left\{ (e^T x)^2 : x_i x_j = 0 \forall ij \in E, \ x^T x = 1, \ x \geq 0 \right\}$$

$$\leq \max_{X \in S_+^{n \times n}} \left\{ ee^T \cdot X : I \cdot X = 1, \ (xx^T)_{ij} = 0 \forall ij \in E, \ X \geq 0 \right\}$$

$$= \max_{X \in S_+^{n \times n}} \left\{ ee^T \cdot XX^T : X \geq 0 \right\}$$

Observations:

- If we include the constraint $X \geq 0$, the number of constraints will be about $n^2$, since we would have a constraint per edge ($x_{ij} = 0 \forall ij \in E$) and one for each non-edge ($x_{ij} \geq 0 \forall ij \notin E$).

- Given an optimal solution $X$ to the problem defining $\vartheta(G)$, we can try to get an approximate solution to $\alpha(G)$ using eigenvalue decomposition.

Before proving the second inequality, we will find the dual of the SDP that defines $\vartheta(G)$:

(P) \quad \max_{X \in S_+^{n \times n}} \quad ee^T \cdot X

$$-I \cdot X = -1,$$

$$(e_ie^T_j + e_je^T_i) \cdot X = 0, \quad ij \in E,$$

$$X \geq 0$$

$$\Updownarrow$$

(P) \quad \min_{X \in S_+^{n \times n}} \quad (-ee^T) \cdot X

$$-I \cdot X = -1,$$

$$(e_ie^T_j + e_je^T_i) \cdot X = 0, \quad ij \in E,$$

$$X \geq 0.$$
The dual of this problem is:

\[
(D) \quad - \max (-I)\eta + \sum_{ij \in E} (e_ie_j^T + e_je_i^T)y_{ij} \leq -ee^T
\]

\[
\Downarrow
\]

\[
(D) \quad \min \eta - ee^T - V \succeq 0,
\]

\[
v_{ij} = 0, \; ij \notin E,
\]

\[
v_{ii} = 0, \; \forall i
\]

\[
\Downarrow
\]

\[
(D) \quad \min \{\lambda_{\max}(V + ee^T) : v_{ij} = 0, \; ij \notin E, \; v_{ii} = 0, \; \forall i\}
\]

By weak duality, \(\vartheta(G)\) is at most the optimal value of (D). In fact, as we will see later, we have equality (strong duality), since both (P) and (D) have strictly feasible solutions (feasible solutions with X or S pd):

- \(\frac{1}{n}I\) is strictly feasible for (P).
- We can get a strictly feasible solution for (D) by taking \(V = 0\), and \(\eta\) as large as we need to make \(\eta I - ee^T \succ 0\).

**Theorem 1** (Lovász’s Sandwich Theorem) \(\alpha(G) \leq \vartheta(G) \leq \chi(G)\).

**Proof:**

*Another proof of left-hand inequality:*

Choose any stable set \(S\) and let its characteristic vector be \(e_S\) (1’s for components \(j\) with \(j \in S\), 0’s ow). Then \(X = \frac{1}{|S|}e_se_s^T\) is feasible in (P), with objective value \(|S|\).

*Proof of right-hand inequality:*

Let \(C_1 \cup C_2 \cup \ldots \cup C_p\) be a clique cover of \(G\). Assume \(C_1 = \{1, \ldots, n_1\}, \; C_2 = \{n_1 + 1, \ldots, n_1 + n_2\}, \ldots, C_p = \{n_1 + \ldots + n_{p-1} + 1, \ldots, n\}\). We want to find a matrix \(V \in S\mathbb{R}^{n \times n}\) with \(v_{ij} = 0\) if \(ij \notin E\), \(v_{ii} = 0\) all \(i\), and \(pI - ee^T - V \succeq 0\). In fact, we choose

\[
pI - ee^T - V = \begin{pmatrix}
p - 1 & \cdots & p - 1 \\
\vdots & \ddots & \vdots \\
p - 1 & \cdots & p - 1
\end{pmatrix}
\]

\[
= \sum_{k<l} (e_{C_k} - e_{C_l})(e_{C_k} - e_{C_l})^T \succeq 0.
\]

\(\square\)
Remarks:

- For perfect graphs (\(\alpha(H) = \chi(H)\)) all vertex-induced subgraphs \(H\) of \(G\), \(\vartheta(G)\) shows that we can compute \(\alpha\) and \(\chi\) for any such graph in polynomial time.

- In general, the gaps between \(\alpha\) and \(\vartheta\) and between \(\vartheta\) and \(\chi\) can be arbitrarily bad! But \(\vartheta\) gives another bound. For example:

\[ s_1 \]
\[ s_2 \]
\[ s_3 \]
\[ s_4 \]
\[ s_5 \]

Let \(s_1, s_2, \ldots, s_5 \in \Sigma\) (alphabet), where \(s_i\) and \(s_j\) can be confused in message iff \(ij \in E\). Then

\[ \alpha(G) = 2 \rightarrow \text{maximum number of non-confoundable messages using 1 symbol (e.g. } s_1, s_3) \]
\[ \alpha(G^2) = 5 \rightarrow \text{maximum number of non-confoundable messages using 2 symbols (e.g. } s_1s_1, s_2s_3, s_3s_5, s_5s_4, s_4s_2) \]

Hence \(\sqrt{\alpha(G^2)} = \sqrt{5}\).

The **Shannon capacity** of a graph is an important measure in information theory and it is given by \(\sup(\alpha(G^k))^{\frac{1}{k}}\). But \(\vartheta(G^k) = (\vartheta(G))^k\), so \(\vartheta(G)\) is an upper bound on the Shannon capacity (e.g. for the pentagon \(\vartheta(C_5) = \sqrt{5}\), and so the Shannon capacity is exactly \(\sqrt{5}\), and this solved a long-standing open question).

**Final remark:**
There are several ways to define \(\vartheta\), e.g.

\[ \vartheta(G) = \min_{\text{all orthon. repr. of } G} \max_{c, u_1, \ldots, u_n} \frac{1}{(c^T u_i)^2} \quad (1/0 := \infty) \]

**Definition 3** \(c, u_1, \ldots, u_n \in \mathbb{R}^n\) form an orthonormal representation of \(G\) if the following holds:

\[ ||c||_2 = ||u_1||_2 = \ldots = ||u_n||_2 = 1 \text{ and } u_i^T u_j = 0 \text{ if } ij \notin E \]

For example, for the pentagon \(C_5\), the optimal orthonormal representation is a five-ribbed umbrella, opened just the right amount, where the \(u_i\)'s are the ribs and \(c\) is the handle.