Lecture 18: Lovasz theta function

Rajat Mittal

IIT Kanpur

We studied the approximation algorithm given by SDP for the problem of max-cut in a graph. Today, the focus will be to look at the independent set of a graph. We will look at the SDP relaxation. It turns out that the relaxation is not tight. Still using properties of SDP, useful information about these quantities can be obtained.

This relaxation was discovered by Lovasz in 1979 for his seminal paper to find out the Shannon capacity of pentagon. The same relaxation has been used in various domains since then. The following notes are made with lot of help from that paper.

1 Independent set and chromatic number of a graph

Given a graph $G = (V, E)$, the independent set of the graph is the subset S of V, s.t., there is no edge between any two vertices of S. Maximum independent set problem is to find out the size of maximum independent set in the given graph. The problem is known to be NP-hard and hence it is not expected to have a polynomial time algorithm.

The complement of the graph $G = (V, E)$ is the graph $\bar{G}$ with the same vertex set and exactly the opposite edge set. So $(i, j) \in E(\bar{G})$ iff $(i, j) \notin E(G)$. The independent set will be a clique (where all vertices are connected) in the complement graph.

A coloring of a graph is a coloring of vertices, s.t., no two adjacent vertices have the same color. In other words, a k-coloring is a mapping $c : [k] \Rightarrow V$, s.t., $c(i) \neq c(j)$ if (i, j) is an edge. The chromatic number of a graph G is the minimum k for which graph G has a k-coloring.

The maximum independent set of graph G is denoted as $\alpha(G)$ and the chromatic number as $\chi(G)$.

2 Lovasz theta number

Lets cast the independent set as an integer program and try to relax it. Again assign a variable y_i to every vertex. The variable takes value 1 if it is part of an independent set and 0 otherwise. Hence the maximum independent set problem can be formulated

$$\begin{aligned} &\max \textstyle\sum_{i\in V} y_i \\ \text{s.t.}\ \ & y_i y_j = 0 \ \ \forall (i,j) \in E \\ & y_i \in \{0,1\} \ \ \forall i \end{aligned} \tag{1}$$

The semidefinite relaxation in this case is not as straightforward as the case of maximum cut.

Exercise 1. Show that the following SDP as relaxation is useless.

$$\begin{aligned} &\max \textstyle\sum_{i\in V} y_i^T y_i \\ \text{s.t.}\ \ & y_i^T y_j = 0 \ \ \forall (i,j) \in E \\ & \|y_i\| = 1 \forall i \in V \end{aligned}$$

We take the following SDP as the relaxed program. Now y_i's are vectors.

$$
\begin{aligned}
&\max \textstyle\sum_{(i,j)\in E} y_i^T y_j \\
\text{s.t.}\quad & y_i^T y_j = 0 \;\; \forall (i,j) \in E \\
&\textstyle\sum_{i\in V} y_i^T y_i = 1
\end{aligned}
\tag{2}
$$

It is almost the vector version of integer program 1. Though we square the optimization function first and then re-normalize the summation of norms to 1 (Eqn. 2).

Exercise 2. Show that any solution of program 1 can be converted into a solution of program 2 without changing the optimization value. Hence program 2 is a relaxation of program 1.

The vector version can be easily converted into matrix version.

$$
\begin{aligned}
&\max \textstyle\sum_{(i,j)\in E} X_{ij} \\
\text{s.t.}\quad & X_{ij} = 0 \;\; \forall (i,j) \in E \\
&\textstyle\sum_{i\in V} Tr(X) = 1 \\
& X \succeq 0
\end{aligned}
\tag{3}
$$

The value of this program 3 is called the Lovasz theta number for graph G ($\theta(G)$). From the fact that it is a relaxation it is clear that $\alpha(G) \leq \theta(G)$. Now we take the dual of this program.

$$
\begin{aligned}
&\min \lambda \\
\text{s.t.}\quad & \lambda I - A \succeq 0 \\
& A_{i,j} = 1 \text{ if } i = j \text{ or } (i,j) \notin E
\end{aligned}
\tag{4}
$$

Exercise 3. Can you give an interpretation of this dual program in terms of eigenvalues of A?

2.1 Orthonormal representation of a graph G

Lets introduce the concept of orthonormal representation of graph to get another definition of Lovasz theta number. Actually Lovasz in his paper used this definition first to come up with other definitions we have mentioned.

An orthonormal representation of a graph G is an assignment of *unit* vector v_i to every vertex i, such that, $v_i^T v_j = 0$ whenever (i,j) is not an edge. The dual program 4 has similar kind of constraints on the entries of matrix A. Lets construct an orthonormal labeling of graph G from the solution of the dual.

Since $\lambda I - A \succeq 0$, its i, j^{th} entry can be written as $y_i^T y_j$ for n vectors $y_1, \cdots, y_n$. Suppose c is some unit vector orthogonal to all vectors $y_1, \cdots, y_n$. Then $v_i = \frac{1}{\sqrt{\lambda}}(c + y_i)$ is an orthonormal labeling of G (prove it as an exercise).

Here $\lambda = (\frac{1}{c^T v_i})^2$. Consider another optimization,

$$
\begin{aligned}
&\min_{c,\{v_i\}} \max_i \left(\tfrac{1}{c^T v_i}\right)^2 \\
\text{s.t.}\quad & \|c\| = \|v_i\| = 1 \;\; \forall i \\
& v_i^T v_j = 0 \;\; \forall (i,j) \notin E
\end{aligned}
\tag{5}
$$

From the above discussion, every solution of the program 4 is also a solution of the program 5 with the same objective value. So the objective value of this new program is less than $\theta(G)$. It turns out that this value is actually equal to $\theta(G)$.

Proof. To prove that $\theta(G)$ is less than the optimal value of program 5. Consider the $c, \{v_i\}$, optimal for the above program. Lets construct matrix A using these vectors.

$$A_{ii} = 1,\ A_{ij} = 1 - \frac{v_i^T v_j}{(c^T v_i)(c^T v_j)}$$

From the construction, the matrix A satisfies the conditions for being the feasible solution of program 4. The aim is to show that $M = (\max_i \frac{1}{c^T v_i})^2 I - A \succeq 0$ (Exercise: Why is that enough).

We will show that matrix $N = D - A$, where D is the diagonal matrix with the i^{th} entry as $(\frac{1}{c^T v_i})^2$ is positive semidefinite. Since M is the addition of N and another diagonal positive semidefinite matrix, it will imply that $M \succeq 0$.

Consider vectors $y_i = c - \frac{v_i}{c^T v_i}$. Then N is the gram matrix of vectors y_i and hence $N \succeq 0$. This completes the proof that optimal value of program 5 is greater than $\theta(G)$. □

Hence the program 5 gives another characterization of Lovasz theta number.

2.2 Chromatic number and Sandwich theorem

Using the new characterization a bound on chromatic number of the complement graph can be obtained. Suppose the chromatic number of the complement graph G is $k = \chi(\bar{G})$ using coloring $C : [n] \to [k]$.

Then consider the vectors $v_i \in \mathbb{R}^k$, where each co-ordinate is assigned to one color. The vector for vertex v_i is e_j (the standard basis vector) if the vertex i is colored with color j. Consider a non-edge (i, j) in G. It is an edge in $\bar{G}$. So i and j are colored differently for the above coloring, and hence $v_i^T v_j = 0$.

Take c to be $\frac{1}{\sqrt{k}} \sum_k e_k$. Then $c, \{v_i\}$ is a feasible solution for program 5. This implies that the k is greater than $\theta(G)$. So we obtain the famous sandwich theorem,

$$\alpha(G) \leq \theta(G) \leq \chi(\bar{G}).$$

Exercise 4. Prove that $\alpha(G) \leq \chi(\bar{G})$ directly (without using $\theta(G)$).