Semidefinite Programming
Lecture 8

scribe: ÜNSAL ÖZDOĞRU

April 3, 1995

COMBINATORIAL OPTIMIZATION

1 introduction

There are two lines of possible applications. One is, suppose that we have a NP-hard Combinatorial Optimization Problem which can not be solved in polynomial time. But we may want to approximate it in polynomial time and find an upper bound (or lower bound). It can be also proved how good these upper or lower bounds are in the case of maximization or minimization problems. Max-Cut problem is one of the application of these. Another approach is basically the following. If we have a general Integer Programming Problem (IPP) it is possible to relax it. We know how to do it in LP. But, what can be said about SDP relaxation. It gives another line of work. The best example of this is Lovàsz' θ function.

We are going to concentrate on a specific problem which is maximum independent set problem for graphs. A maximum independent set for a graph is a subset of mutually nonadjacent vertices of maximum cardinality. More generally, one can assign weigths on the vertices of a graph. Problem becomes a weighted max independent set problem in which we seek for the heaviest maximum independent set where the weight of a subgraph is defined as the sum of its weights. That is more general case. It is easy to write this problem as an IPP.

$$\begin{array}{lll} \mathbf{IP}: & \max & \mathbf{w}^T\mathbf{x} \\ & & x_i + x_j \leq 1 \quad \forall i,j \notin E \\ & & \\ & & x_i \in \{0,1\} \end{array}$$

This is a maximum independent set problem. In this formulation any solution $\mathbf{x}$ is the characteristic vector of an independent set. Suppose that we've a general set :

$$\mathbf{x} = (1, 2, \ldots, n)$$

Now, take a subset of $\mathbf{x}$. For instance, $\mathbf{y} = (2,5,9) \subseteq \mathbf{x}$. This subset can be represented by an n-vector. Namely, the second, fifth and nineth entries of the vector are 1 , anything else is 0.

$$(0,1,0,0,1,0,0,0,1,0,\ldots,0)$$

is called the characteristic vector of set $\mathbf{y}$. This statement says, if edges i and j are in the set, x_i and x_j both can not be 1. Here, number of constraints are precisely the number of edges in the graph.

This problem is a well research NP-hard problem solved. One way of solving is to drop constraints, in fact relaxation does mean dropping of some constraints.

$$\begin{aligned} \mathbf{LP}: \quad & \max \mathbf{w}^T\mathbf{x} \\ & x_i + x_j \leq 1 \\ & 0 \leq x_i \leq 1 \end{aligned}$$

For what class of a graph, relaxation does not affect the problem? These are bipartite graphs. It can be written as a matrix form, such that,

$$\mathbf{Ax} \leq 1$$

where $\mathbf{A}$ is a incidence matrix that every column of it precisely includes two $1's$.

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & \ldots & \ldots & 0 \\ 0 & 1 & \ldots & \ldots & 1 \\ 1 & 0 & \ldots & \ldots & 1 \\ 0 & 0 & \ldots & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 1 & \ldots & \ldots & 0 \end{pmatrix}$$

This matrix is totally unimodular (TU) in which the determinant of every square submatrix either $-1, 0, 1$. If matrix of a bipartite graph is TU then, its solutions will be integral.

THEOREM 1.1. *All solutions of LP are integral iff graph $G = (V,E)$ is bipartite.*

Proof. Part A: Bipartite $\Rightarrow$ TU. Proof is by induction. Our goal is to show that if $\mathbf{A}$ is coming from a bipartite graph, all of the kxk submatrices will be determinants $0, -1, 1$. Now, assume that $k = 1$. It is obvious that all 1x1 matrices will be $0, -1, 1$. Now again, assume that rxr submatrices of kxk matrices determinat are $0, -1, 1$. Here, $r < k$. Now, we want to prove that all kxk submatrices will be determinant $0, -1, 1$. There are three cases :

1. Case: If there is a column of all zeros: Clearly, it means

$$\Rightarrow det = 0$$

. We are done.

2. Case: If there exits a column with exactly 1: Expand it and by induction,

$$\Rightarrow det = 0, -1, 1$$

   We re done again.

3. Case: If every column has exactly two 1's. Since graph is bipartite, it can be divided into two parts. Say, some edges go to the one part of it, rest of them go to the other part. Then, rows of the graph can be rearranged. By the rearranging of the rows, the graph will be in the different form, such that on each column will have only one 1 among in the first t rows and one 1 among the last $n-(t+1)$ rows. Then, determinat will be zero, because the number of the ones above the first t rows will be equal to the number of the ones below the last $t+1$ rows since there is $1-1$ correspondence between each column. There is a linear independence among the rows of a matrix. Therefore,

$$\sum_{i=1}^{t} row_i A = \sum_{i=t+1}^{n} row_i A$$

   Therefore, Det $\mathbf{A} = 0$.

Part B: TU $\Rightarrow$ Bipartite.

Observation Det $(A_c) = \mp 2$ for C which is an odd cycle. If C is a cycle with odd number of vertices, determinant of its incidence matrix ∓ 2. Proof is exercise. That is also essentially proof of the theorem. Because, if graph is bipartite iff it has no odd cycle. If it has an odd cycle, A can not be TU. ■

Feasible set of LP is a polytope which is called STAB(G),

STAB(G) = Conv(x: x is a characteristic vector of some stable set in G).

Every vertex of STAB(G) polytope is a feasible for LP and every convex combination of this is also feasible.

Some classes of valid inequalities for STAB(G)

1. Clique Constraint: Suppose k is a clique. What is the intersection of an arbitrary clique and an arbitrary stable set?

$$|K \bigcap S| \leq 1$$

   It can be written as the following form :

$$1_K x \leq 1$$

   where 1_K is a characteristic vector of a given clique and inequality is satisfied by every vertex of STAB(G). Maximal cliques which are not proper subset of any other clique.

$$K_1 \subseteq K_2$$
$$1_{K_1}\mathbf{x} \leq 1 \Leftarrow 1_{K_2}\mathbf{x} \leq 1$$

Let us define following polyhedran :

$$Q - STAB(G) = \{x : 1_K^T \leq 1,\ x_i \geq 0,\ \forall k \in G\}$$

2. Odd Hole Constraints: C is an odd cycle. $|C| = 2k + 1$ is number of vertices. Now, take an arbitrary stable set. What is the intersection of an arbitrary stable set and this odd cycle. That is,

$$|C \cap S| < k$$

The max independent set of an odd cycle is always k if it has $2k + 1$ vertices. So, the corresponding inequalities are as the following,

$$C - STAB(G) = \{x : 1_C^T x \leq k, \forall\ cycle\ C\ where\ |C| = 2k + 1, 0 \leq x_i \leq 1\}$$
$$STAB(G) \subseteq C - STAB(G)$$

Clique Constraints are Perfect Graphs. STAB(G) is an intersection of Q-STAB(G) and C-STAB(G).

3. An Odd Anti-hole constraints: Complements of cycles. Suppose that we are taken a clique and removed one odd cycle. The left is anti-hole. The cardinality of intersection of an anti-hole and a stable set is,

$$|\bar{C} \cap S| \leq 2$$

$$\bar{C} - STAB(G) = \{x : 1_C^T x \leq 2, \forall\ \bar{C}\ anti-holes, x_i \geq 0\}$$

Let's take a special class of a graph for which we know the maximum independent set of the problem. Then, if there is such a subgraph, then intersection of maximum independent set and $\bar{C} - STAB(G)$ can not be larger. Suppose that, such a class be a cycle and we've a formula to compute maximum independent set of the problem of a particular subgraph. Then, inequalities can be always created. If G was such a class, then

$$1_G^T x \leq \alpha(G)$$

where $\alpha(G)$ is a maximum independent set of G. The intersection of n independent set and such a graph cannot be more than $\alpha(G)$ and it is easily computable if G is bipartite.

4. Odd wheel inequalities: Suppose that, wheel has $1, 2, \ldots, 2k - 1$ vertices and the vertex 2k is in the center.

Claim is that following inequality is valid:

$$\sum_{i=1}^{2k-1} x_i + (k - 1)x_{2k} \leq k - 1$$

First and second term of the left handside are mutually exclusive. If independent set intersects with the vertices, either it picks the first term (it can not pick another vertices) or vice versa.

The maximum independent set of

$$\sum_{i=1}^{2k-1} x_i \leq k-1$$

is $k-1$.

$$|W \cap S| < k-1$$

$$\sum_{i=1}^{2k-1} x_i \leq k-1$$

$$(k-1)x_{2k} \leq k-1$$

$$\textbf{either} \quad \textstyle\sum_{i=1}^{2k-1} x_i = 0 \ or \ (k-1)x_k \leq 1 \quad \textbf{or} \quad \textstyle\sum_{i=1} 2k-1x_i = 1 \ or \ (k-1)x_k = 0$$

Then,

$$W - STAB(G) = \{x : \textstyle\sum_{i \in P(W)} x_i + (k-1)x_{2k} \leq (k-1), \ x_i geq0\}$$

for a wheel in G. That also defines a polytope.

DEFINITION 1.1. *A graph is perfect graph if*

$$STAB(G) = Q - STAB(G)$$

Lifting Think of a polytope in $N-$dimensional space and m consraints (facets).

Suppose that we project a polytope into lower dimensional space, now we've n dimensional space and M-facets. Projecting of a polytope is still another polytope, but it will have different number of facets. Now, we know that $n < N$. Projecting into a lower dimensional space sounds some information is losing, but meanwhile many facets are gained.

Let's look at another point of view. There is a polytope in n-dimensional space with m-facets. During the lifting process, new variables are added which are expessed as redundant variables.

$$N - dim \overset{\textbf{projection}}{\Rightarrow} m - facets$$

$$n - dim \overset{\textbf{lifting}}{\Rightarrow} M - facets$$

2 An Example

Parity Polytope

$$\mathbf{PP} = Conv\{x \in \{0,1\}^n : there\ are\ odd\ number\ of\ ones \in \mathbf{x}\}$$

Suppose that, we want to optimize over that polytope. There is a vector w, and we want to maximize it over the polytope. We can formulate it as a following format.

$$\begin{array}{rl} \max & \mathbf{w}^T x \\ \text{st} & \mathbf{x} \in PP \end{array}$$

if n is odd then we take the vector of all ones, if n is even remove the smallest one. Optimization over parity polytope is a trivial problem and this polytope has exponential number of inequalities (or facets). Large number of vertices does not mean large number of inequalities. Counter example is cube. It has 2^n vertices but it can be represented by $2n$ inequalities.

It works as follows : every point inside in a small polytope, there exist an odd number number k, such thatx can be written as,

$$\begin{array}{rl} \mathbf{k} : \mathbf{odd} & \sum_{i=1}^{n} \alpha_i y_i = x \\ & \sum_{j=1}^{n} (z_i)_j = k\alpha_i,\ \forall i \\ & \sum_{i=1}^{k} (z_i) = \mathbf{x},\ \forall i \\ & \sum_i \alpha_i = 1 \\ & 0 \le x_i \le 1 \end{array}$$

z's, x's, α's are all polynomially many variables. But number of inequalities is small. It was a lifting process, because, now we've z's, α's in additional to x. Number of variables increases polynomially, after adding redundancy, number of inequalities is also polynomially. Because if x is inside polytope then it can be represented as a convex combination of y_i's which are the vertices. The argument key word here, every point inside the polytope is a convex hull of a bunch of $\{0,1\}$ vectors with odd number of variables.

Now, we have basically $n^2 + n + n$ number of variables.

One question can be asked? Why do we do lifting process? There is a special class of graphs which is Perfect graphs we will concentrate on this. It turns out that in the class of perfect graphs we can compute maximal independent set problem and also maximal clique problem in polynomial time. the trick is based on lifting. the clique polytope Q-STAB(G) has exponential number of inequalities in SDP. So there exists a SDP whose projection is some direction which is precisely Q-STAB(G) in the case of perfect graphs. So, on the other hand, nobody knows a small linear programming expression of Q-STAB(G) by lifting. It is possible that ther is a lifting of Q-STAB(G) for perfect graphs into another bigger LP, but still it has polynomial number of inequalities. There is a sub exponential lifting $nlogn$ inequalities, but not polynomially.

Definition 2.1. *In G= (V, E) an orthonormal represantation of G is a set of vectors* $u_1, u_2, \ldots, u_n$ *with the following properties:*

$$\mathbf{1}.\|u_i\| = 1$$

$$\mathbf{2}.u_i \perp u_j(ie, u_i^T u_j = 0)\ \forall\, i, j \notin E$$

Each vertex associate with a bunch of 1 vector. They are unit length and if i and j are not adjacent to each other, their corresponding vector should be orthogonal to each other. Obviously, every graph has infinitly many orthonormal represantations.

observation 1: Let C be another vector with unit vector. Claim :

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

where S is a stable set. Define the following object :

$$TH(G) = \{x \in \mathbf{R} : \sum_{i=1}^{n}(C^T u_i)^2 x_i \leq 1, \|k\| = 1\}$$

where $\forall u_i$ orthonormal representation of G.

If we take arbitrary orthonormal represantation of G and arbitrary C, we will get infine number of inequalities which are halfspaces. Set is convex. This observation shows that if we put x, the characeristic vector of an independent set is always satisfied.

$$STAB(G) \subseteq TH(G)$$

$C^T u_i$ is the cosine of the angle between C and each u_i, basically. This is essentially $(cosine)^2$ of C and each u_i. If we rotate the space, u_i's become $e_1, e_2, \ldots, e_n$ standart basis. This is nothing, but ith coordinate of C. Therefore, sum of the ith coordinate of C in the space is 1. TH(G) is true for any stable set. We can also conclude

$$Q - STAB(G)\ is\ a\ subset\ of\ TH(G)$$

Observation 2: Suppose k is a clique in our graph.

$$\begin{aligned} &\|u_i\| = 1, \quad \forall i \notin K, \quad u_i \perp u_j, \quad i, j \notin K \\ &\|C\| = 1, \quad C \perp u_i, \quad i \notin K \end{aligned}$$

By definition, this is all of the inequalities arising from the orthonormal representation.

$$\sum_{i=1}^{n}(C^T u_i)^2 x_i \leq 1$$

$$\sum_{i \in K} x_i \leq 1$$

$$STAB(G) \subseteq TH(G) \subseteq Q - STAB(G)$$

in the case of perfect graph. Let's define three numbers:

$$\begin{array}{ll}
\alpha(G,W) = \max \mathbf{w}^T \mathbf{x} & \\
\text{st} & \mathbf{x} \in STAB(G) \\
\theta(G,W) = \max \mathbf{w}^T \mathbf{x} & \\
\text{st} & \mathbf{x} \in TH(G) \\
\alpha^*(G,W) = \max \mathbf{w}^T \mathbf{x} & \\
\text{st} & \mathbf{x} \in Q - STAB(G)
\end{array}$$

That inclusuion gives:

$$\alpha(G,W) \leq \theta(G,W) \leq \alpha^*(G,W)$$