MIT 6.854/18.415: Advanced Algorithms	Spring 2016
Lecture 20 – April 20, 2016	
Prof. Ankur Moitra	*Scribe: Dhiraj Holden*

In this lecture, we will talk about more sophisticated ways of rounding SDPs which work when the plain hyperplane rounding fails. In particular, we will prove Grothendieck's inequality using by rounding an SDP with a modified version of hyperplane rounding. We will also introduce the Lovasz theta function, which serves as both an upper bound on the largest independent set of the graph and a lower bound on the chromatic number of the complement of a graph, and can be computed efficiently using SDPs.

1 Grothendieck's Inequality

Consider the quadratic program $\max_{x_i,y_j\in\{\pm 1\}}\sum_{i,j}A_{ij}x_iy_j$. We think of A as a $m\times n$ matrix. MAXCUT is a special case of this problem, if we set $A_{ii}=2|E|$ for all i to force $x_i=y_i$ and $A_{ij}=-1/4$ if $i,j\in E$ and 0 otherwise, as then what we are maximizing is $\sum_{i,j\in E}1/2-\frac{x_ix_j}{2}$. We will look at a very powerful SDP relaxation to this problem, which will be different than the randomized rounding algorithm for MAXCUT. To formulate the relaxation, we proceed like we did with MAXCUT, by replacing x_i, y_j with vectors and approximating with SDPs. The SDP relaxation we will consider is equation (*) below.

(*) $\max\sum_{ij}A_{ij}Z_{ij}$, where Z is the upper right block of an $(m+n)\times(m+n)$ block matrix B with 1's on the diagonal.

We will proceed to show how this gives us vectors representing the x_i and y_j. B is $(m+n)\times(m+n)$, and $B=WW^T$ as it is PSD. We can think of splitting W into the first m rows and the last n rows, and we can think of the first m rows of W as unit vectors $u_1,...,u_m$ and the last n rows as vectors $v_1,...,v_n$. We can rewrite the problem as $\max_{\|u_i\|=1,\|v_j\|=1}\sum_{i,j}A_{ij}\langle u_i,v_j\rangle$ as $z_{ij}=\langle u_i,v_j\rangle$. This is a relaxation of the original problem because we can set $u_i=[x_i,0,...,0]$, $v_j=[y_j,0,...,0]$ where x_i,y_j are the solution to $\max_{x_i,y_j\in\{\pm 1\}}\sum_{i,j}A_{ij}x_iy_j$ and this is a feasible solution to the SDP.

Let OPT be the solution to the original problem and let OPT' be the solution of the SDP.

Theorem 1 (Grothendieck's inequality [2, 3])**.** *$OPT \le OPT' \le \frac{\pi}{2\ln(1+\sqrt{2})}OPT$.*

For a long time this was conjectured to be the right answer, until 3 years ago [1] showed that you could subtract a constant on the order of 10^{-10}. In fact, people have given algorithms that within a finite amount of time would calculate Grothendieck's constant to finite precision.

What makes this problem different from MAXCUT? Can we just use hyperplane rounding, and set $x_i=sgn(a^Tu_i)$, $y_j=sgn(b^Tv_j)$? The way our analysis worked last time was that we looked at the expected contribution to the SDP versus the expected contribution of the rounding. However, this will not work in this case as it could be the case that the contribution to the positive terms becomes much less than the contribution to the negative terms because of the nonlinear relationship. $A_{ij}\langle u_i,v_j\rangle$ is the contribution to the SDP, and the contribution to the expected value of hyperplane rounding is equal to $A_{ij}(\frac{-\theta_{ij}}{\pi}+(1-\frac{\theta_{ij}}{\pi}))=A_{ij}\frac{2}{\pi}(\frac{\pi}{2}-\theta_{ij})=A_{ij}\frac{2}{\pi}\arcsin(\langle u_i,v_j\rangle)$. What we're worried about is the case when SDP gives us +7 -6 +7 -6 and the value of the rounding is $+\alpha 7$ -6 $+\alpha 7$ -6, which gives us a value very close to 0.

Fact: $\mathbb{E}[sgn(a^T u_i)sgn(a^T v_j) = \frac{2}{\pi}\arcsin(\langle u_i, v_j\rangle)$

To fix this, we will create new vectors such that a linear relationship holds, allowing us to use hyperplane rounding and get a reliable approximation.

Lemma 2. *For any unit vectors $u_1, ..., u_m, v_1, ..., v_n$, there is a new set of unit vectors $u'_1, ..., u'_m, v'_1, ..., v'_n$ with the property that $\mathbb{E}[sgn(a^T u'_i)sgn(a^T v'_j)] = \frac{2}{\pi}\ln(1+\sqrt{2})\langle u_i, v_j\rangle$.*

Definition 3. $u^{\otimes 2} = [(u_1)^2, u_1u_2, u_1u_3, ...]$*, and similarly for* $u^{\otimes n}$*.*

Proof. Guess a constant $c = \sinh^{-1} 1 = \ln(1+\sqrt{2})$. Then by Taylor's theorem $\sin(c\langle u, v\rangle) = \sum_{k=0}^{\infty}(-1)^k \frac{c^{2k+1}}{(2k+1)!}\langle u, v\rangle^{2k+1}$. From the Taylor series we can get the infinite-dimensional vectors by inspection. Since $\langle u, v\rangle^{2k+1} = \langle u^{\otimes 2k+1} v^{\otimes 2k+1}\rangle$, we can set $u' = [\sqrt{c}u, (-1)\sqrt{\frac{c^3}{3!}}u^{\otimes 3}, ...(-1)^k\sqrt{\frac{c^{2k+1}}{(2k+1)!}}u^{\otimes 2k+1}...]$ and $v' = [\sqrt{c}v, \sqrt{\frac{c^3}{3!}}v^{\otimes 3}, ...]$. Then $\langle u', v'\rangle = \sin(c\langle u, v\rangle)$ as expanding the inner product gives us the Taylor series. Moreover, $\|u'\|^2 = \sinh(c\|u\|^2) = \sinh(sinh^{-1}(1)) = 1$, and $\|v'\|^2 = 1$ as well. Finally $\mathbb{E}[sgn(a^T u')sqn(a^T v')] = \frac{2}{\pi}\arcsin(\langle u', v'\rangle) = \frac{2}{\pi}\arcsin(\sin(c\langle u, v\rangle)) = \frac{2}{\pi}c\langle u, v\rangle$. □

The catch is that the u', v' are infinite-dimensional, and so care must be taken to make sure working with u', v' can be done efficiently. The lemma actually implies 1.

Proof. Using hyperplane rounding, find u_i, v_j with $\langle u_i, v_j\rangle = z_{ij}$ and then use the procedure in 2 to get u'_i, v'_j. Then, round the u'_i and v'_j by choosing a uniformly at random and then set $x_i = a^T u'_i$ and $y_j = a^T v'_j$. To analyze this procedure, consider some term i, j which has $A_{ij}\langle u_i, v_j\rangle$ as its contribution to the SDP and $A_{ij}\frac{2}{\pi}\ln(1+\sqrt{2})A_{ij}\langle u_i, v_j\rangle$ as its contribution to the expected value. This gives us Grothendieck's inequality. □

2 Lovasz Theta Function

Let $\alpha(G)$ be the size of the largest independent set. $\bar{\chi}(G)$ is the chromatic number of $\bar{G}$ (the complement of G). Then $\bar{\chi}(G) \geq \alpha(G)$, as the size of the largest independent set in the graph is the size of the largest clique in its complement.

Definition 4. $\theta(G)$ *is defined as* $\min k$ *s.t.* $\langle v_i, v_j\rangle = \frac{-1}{k-1}\forall(i, j) \notin E$ *and* $\langle v_i, v_i\rangle = 1$.

Theorem 5. $\alpha(G) \leq \theta(G) \leq \bar{\chi}(G)$*.*

Proof. $\theta(G) \leq \bar{\chi}(G)$: I claim there exist k unit vectors $u_1, ..., u_k$ where $\langle u_i, u_j\rangle = \frac{-1}{k-1}$ and $\langle u_i, u_i\rangle = 1$ using the vectors from the centroid of the simplex to each corner. Thus any k-coloring of $\bar{G}$ yields a feasible solution when we assign one of the unit vectors to the vertices of each color.

$\alpha(G) \leq \theta(G)$: Take some optimal solution to the SDP and takes vectors $v_1, ..., v_S$ of an independent set. Then $0 \leq (\sum_{i=1}^{S} v_i)^T(\sum_{i=1}^{S} v_i) = s + \sum_{i\neq j}\langle v_i, v_j\rangle$. By an averaging argument, there exists $i \neq j$ $\langle v_i, v_j\rangle \geq \frac{-s}{2\binom{s}{2}} = \frac{-1}{s-1}$. Now, $\langle v_i, v_j\rangle = \frac{-1}{\theta(G)-1} \leq \frac{-1}{s-1}$, or $s \leq \theta(G)$. □

There are important graphs where $\alpha(G) = \bar{\chi}(G)$ (e.g. perfect graphs). Thus for perfect graphs we can compute the size of the largest independent set efficiently.

References

[1] Braverman, Mark; Makarychev, Konstantin; Makarychev, Yury; Naor, Assaf (2011), "The Grothendieck Constant is Strictly Smaller than Krivine's Bound", *52nd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pp. 453462

[2] Grothendieck, Alexander (1953), "Rsum de la thorie mtrique des produits tensoriels topologiques", *Bol. Soc. Mat. Sao Paulo* 8: 179

[3] Krivine, J.-L. (1979), "Constantes de Grothendieck et fonctions de type positif sur les sphres", *Advances in Mathematics* 31 (1): 1630