Handout 5s
September 7, 2017

Summary of Lecture 5

In which we study the SDP relaxation of Max Cut in random graphs.

1 A Quick Review of Chernoff Bounds

If $X_1, \ldots, X_n$ are mutually independent random variables taking values in $\{0,1\}$, and $X := \sum_{i=1}^n X_i$ is their sum, the following three versions of the Chernoff bounds hold:

$$\begin{aligned} rl\forall 0 \le \epsilon \le 1. \qquad & \mathbb{P}[|X - \mathbb{E}\, X| \ge \epsilon \cdot \mathbb{E}\, X] \le e^{-\Omega(\epsilon^2 \cdot \mathbb{E}\, X)} \\ \forall 0 \le \epsilon \le 1. \qquad & \mathbb{P}[|X - \mathbb{E}\, X| \ge \epsilon \cdot n] \le e^{-\Omega(\epsilon^2 \cdot n)} \\ \forall c > 1. \qquad & \mathbb{P}[|X - \mathbb{E}\, X| \ge c \cdot \mathbb{E}\, X] \le e^{-\Omega((c \log c) \cdot \mathbb{E}\, X)} \end{aligned}$$

2 Cutting a Near-Optimal Number of Edges in $G_{n,p}$ Via SDP Rounding

Consider the $G_{n,p}$ distribution where the average degree $d = p \cdot (n-1)$ satisfies $\log n \le d \le o(n^{1/3})$.

First we argue, using the $d \ge \log n$ assumption, that there is a constant $c > 0$ such that with high probability every node has degree $\le cd$. Indeed, the probability that there exists a node of degree $> cd$ is at most n times the probability that a fixed node v has degree more than cd. If we apply the Chernoff of the third type above to the number of neighbors of v we have

$$\mathbb{P}[\exists \text{ a node of degree } > cd] \le n \cdot e^{-\Omega(d \cdot (c-1) \log(c-1))} \le n \cdot e^{-2d} \le \frac{1}{n}$$

which is valid for a sufficiently large choice of c.

Next we compute the number of nodes that participate in a triangle. A node has probability at most $p^2 n^2$ of participating in a triangle, and so the expected number of nodes participating in triangles is at most $p^2 n^3 = o(n)$. Thus, by using Markov's inequality, there is a $1 - o(1)$ probability that at most, say, $\frac{n}{10c}$ nodes participate in triangles. (Now we have used the assumption that $d = o(n^{1/3})$.

Overall, we have a $1 - o(1)$ probability that every node has degree $\leq cd$, and that at most $dn/10$ edges are incident on nodes that participate in a triangle.

Now, construct an SDP solution as we did last time for all the nodes that do not participate in triangles, and make the nodes that participate in triangles orthogonal to all other nodes.

If we round this solution with a random hyperplane, every edge has probability $\frac{1}{2} + \frac{1}{O(\sqrt{d})}$ of being cut, except that $\leq dn/10$ edges incident on nodes that participate in triangles, which have probability $1/2$ of being cut. Overall, we cut on average $|E|/2 + \Omega(\sqrt{d}n)$ edges, which the asymptotically tight bound that we had achieved with a greedy algorithm, but that now we are able to achieve via a much more robust algorithm, that performs non-trivially well on all graphs.

3 A Quick Review of Eigenvalue Computations as SDPs

Now we have shown that the $|E|/2 + \Omega(\sqrt{d}n)$ lower bound on the probable optimum can be achieved constructively by reasoning about the SDP, we want to show that the $|E|/2 + O(\sqrt{d}n)$ upper bound certificate that we derived from spectral methods can also be obtained through the SDP.

As a starting point, we will show that computing the largest eigenvalue of a symmetric matrix can be formulated as an SDP. Recall that if M is a symmetric matrix then its largest eigenvalue is the solution of the optimization problem

$$\max_{x_1,\ldots,x_n} \frac{\sum_{i,j} M_{i,j} x_i x_i}{\sum_i x_i^2} \tag{1}$$

which we can also write as the homogeneous quadratic optimization problem

$$\begin{array}{ll} \max & \sum_{i,j} M_{i,j} x_i x_i \\ s.t. & \\ & \sum_i x_i^2 = 1 \end{array}$$

We notes in past lectures that once we have a quadratic optimization problem we can always relax it in a automatic way to a SDP. If we do that, we get

$$\begin{array}{ll} \max & \sum_{i,j} M_{i,j} \langle \mathbf{x}_i, \mathbf{x}_i \rangle \\ s.t. & \\ & \sum_i ||\mathbf{x}_i||^2 = 1 \end{array}$$

which we can also write as

$$\max_{\mathbf{x}_1,\ldots,\mathbf{x}_n} \frac{\sum_{i,j} M_{i,j} \langle \mathbf{x}_i, \mathbf{x}_i \rangle}{\sum_i ||\mathbf{x}_i||^2} \tag{2}$$

Although this generally does not happen, we will show that in this case the SDP relaxation is without loss, and the optimization problems (1) and (2) are actually equivalent.

One direction is trivial: every solution to (1) is also a solution to (2) and so the optimum of (1) is smaller than or equal to the optimum of (2).

For the other direction, let c be the optimum of (2) and $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be an optimum solution. We write $\mathbf{x}_{i,k}$ for the k-th coordinate of the vector $\mathbf{x}_i$. Then

$$\begin{aligned}
c &= \frac{\sum_{i,j} M_{i,j} \langle \mathbf{x}_i, \mathbf{x}_i \rangle}{\sum_i ||\mathbf{x}_i||^2} \\
&= \frac{\sum_{i,j} M_{i,j} \sum_k \mathbf{x}_{i,k}, \mathbf{x}_{i,k}}{\sum_i \sum_k \mathbf{x}_{i,k}^2} \\
&= \frac{\sum_k \left(\sum_{i,j} M_{i,j} \mathbf{x}_{i,k}, \mathbf{x}_{i,k} \right)}{\sum_k \left(\sum_i \mathbf{x}_{i,k}^2 \right)} \\
&\leq \max_k \frac{\sum_{i,j} M_{i,j} \mathbf{x}_{i,k}, \mathbf{x}_{i,k}}{\sum_i \mathbf{x}_{i,k}^2} \\
&\leq \max_{x_1, \ldots, x_n} \frac{\sum_{i,j} M_{i,j} x_i x_i}{\sum_i x_i^2}
\end{aligned}$$

where the second-to-last step uses the inequality

$$\frac{\sum_k a_k}{\sum_k b_k} \leq \max_k \frac{a_k}{b_k}$$

which is valid whenever the b_k are positive, because we have

$$\sum_k a_k = \sum_k b_k \cdot \frac{a_k}{b_k} \leq \left(\max_i \cdot \frac{a_i}{b_i} \right) \cdot \sum_k b_k$$

4 The Spectral Norm Certificate as a SDP Certificate

Let's start from the SDP relaxation of Max Cut on a graph G

$$\begin{array}{ll}
\max & \displaystyle\sum_{(i,j) \in E} \frac{1}{4} ||\mathbf{x}_i - \mathbf{x}_j||^2 \\
s.t. & \\
 & ||x_i||^2 \quad \forall i \in V
\end{array}$$

We will refer to the value of the optimum as $SDPMaxCut(G)$. Cleary we have $MaxCut(G) \leq SDPMaxCut(G)$. The following is an equivalent formulation, because, under the constraints, $\frac{1}{4}||\mathbf{x}_i - \mathbf{x}_j||^2 = \frac{1}{2} - \frac{1}{2}\langle \mathbf{x}_i, \mathbf{x}_j \rangle$.

$$
\begin{aligned}
\max \quad & \frac{|E|}{2} - \sum_{(i,j)\in E} \frac{1}{2}\langle \mathbf{x}_i, \mathbf{x}_j\rangle \\
s.t. \quad & \\
& ||\mathbf{x}_i||^2 \quad \forall i \in V
\end{aligned}
$$

Now we relax the constraints $\forall i \in V. ||\mathbf{x}_i||^2 = 1$ to the single constraint $\sum_i ||\mathbf{x}_i||^2 = n$ so that we get the new relaxation $SDP'MaxCut(G)$

$$
\begin{aligned}
\max \quad & \frac{|E|}{2} - \frac{1}{2} \sum_{(i,j)\in E} \langle \mathbf{x}_i, \mathbf{x}_j\rangle \\
s.t. \quad & \\
& \sum_{i\in V} ||\mathbf{x}_i||^2 = n
\end{aligned}
$$

And we clearly have $MaxCut(G) \leq SDPMaxCut(G) \leq SDP'MaxCut(G)$. But we can rewrite $SDP'MaxCut(G)$ as

$$
\frac{|E|}{2} + \frac{n}{4} \cdot \max \frac{\sum_{i,j} -A_{i,j}\langle \mathbf{x}_i, \mathbf{x}_j\rangle}{\sum_{i\in V} ||\mathbf{x}_i||^2}
$$

And from what we have seen in the previous section, this is equivalent to

$$
\frac{|E|}{2} + \frac{n}{4} \cdot \lambda_{\max}(-A)
$$

where we use the notation $\lambda_{\max}(M)$ to represent the largest eigenvalue of M.

Finally, for every $p \geq 0$, we have

$$
\begin{aligned}
SDP'MaxCut(G) &= \frac{|E|}{2} + \frac{n}{4} \cdot \lambda_{\max}(-A) \\
&\leq \frac{|E|}{2} + \frac{n}{4} \cdot \lambda_{\max}(pJ - A) \\
&\leq \frac{|E|}{2} + \frac{n}{4} \cdot \lambda_{\max}||pJ - A||
\end{aligned}
$$

Now, we have claimed that, for $p > \frac{\log n}{n}$, we that the adjacency matrix A of a graph sampled from $G_{n,p}$ satisfies with high probability $||pJ - A|| \leq O(\sqrt{np})$, and so we have that with high probabilty we also have $SDPMaxCut(G) \leq |E|/2 + O(n \cdot \sqrt{d})$, meaning that the polynomial time computable upper bound to $MaxCut(G)$ given by $SDPMaxCut(G)$ has a high probability of being the correct asymptotic bound.