Problem Set 3 Solutions

Samuel B. Hopkins

December 5, 2024

Problem 1 (SoS proof for clique size bound)**.** Let G be a graph drawn from $G(n,1/2)$. Show that with high probability, there exists a sum-of-squares proof of constant degree that certifies that G does not contain a clique of size greater than $O(\sqrt{n\log n})$.

Solution Given a graph G, consider the system of polynomials

$$\mathcal{P} = \begin{Bmatrix} x_i^2 = x_i & \forall i \in [n], \\ x_i x_j = 0 & \forall ij \notin E(G) \end{Bmatrix}.$$

The first collection of constraints indicates that the x_i behave like boolean variables, while the second collection of constraints implies that the set $\{i : x_i = 1\}$ forms a clique. Given a solution to the above system, the size of the corresponding clique is $\sum_{i\in[n]} x_i$.

We shall show that with high probability over $G \sim G(n,1/2)$, $\mathcal{P} \vdash_4 \sum x_i \le O(\sqrt{n\log n})$. Let A be the adjacency matrix of G, and J the all-ones matrix.

Let $\widetilde{\mathbb{E}}$ be an arbitrary pseudoexpectation satisfying $\mathcal{P}$. For starters, we have

$$\left(\widetilde{\mathbb{E}} \sum x_i\right)^2 \le \widetilde{\mathbb{E}} \left(\sum x_i\right)^2 = \widetilde{\mathbb{E}}\, x^\top J x = \widetilde{\mathbb{E}} \sum_{i,j\in[n]} x_i x_j = 2 \cdot \widetilde{\mathbb{E}} \sum_{ij\in E} x_i x_j = 2 \cdot \widetilde{\mathbb{E}}\, x^\top A x.$$

Here, the first inequality is Cauchy-Schwarz, and the second-to-last equality is because $\widetilde{\mathbb{E}}\, x_i x_j = 0$ for any non-edge ij. Now, we have

$$\widetilde{\mathbb{E}}\, x^\top A x = \widetilde{\mathbb{E}}\, x^\top \left(\frac{1}{2}J\right) x + \widetilde{\mathbb{E}}\, x^\top \left(A - \frac{1}{2}J\right) x = \frac{1}{2}\widetilde{\mathbb{E}}\, x^\top A x + \widetilde{\mathbb{E}}\, x^\top \left(A - \frac{1}{2}J\right) x,$$

so

$$\left(\widetilde{\mathbb{E}} \sum x_i\right)^2 \lesssim \widetilde{\mathbb{E}}\, x^\top A x \lesssim \widetilde{\mathbb{E}}\, x^\top \left(A - \frac{1}{2}J\right) x. \tag{1}$$

To conclude, we shall show that with high probability, $\left\|A - \frac{1}{2}J\right\|_{\mathrm{op}} = O(\sqrt{n\log n})$. Given this, we are done: as we saw early in the course, this would imply that $\widetilde{\mathbb{E}}\, x^\top \left(A - \frac{1}{2}J\right) x \le \widetilde{\mathbb{E}} \left\|A - \frac{1}{2}J\right\|_{\mathrm{op}} \sum x_i^2 = O(\sqrt{n\log n}) \cdot \widetilde{\mathbb{E}} \sum x_i$, and plugging this back into (1) completes the proof.

It remains to prove the high-probability bound on the operator norm. Let $B = A - \frac{1}{2}J + \frac{1}{2}\mathrm{Id}$. For each pair of distinct indices $i,j \in [n]$, let $B^{(ij)}$ be the matrix such that $B^{(ij)}_{ij} = B^{(ij)}_{ji}$ are uniformly randomly drawn from $\{\pm 1\}$, and all other entries are 0. Note that B has the same distribution as $\sum_{i,j\in[n]\text{ distinct}} B^{(ij)}$. Clearly, the operator norm of any $B^{(ij)}$ is almost surely bounded by 2. It is also not difficult to check that for any i, j, $\mathbb{E}(B^{(ij)})^2 = e_i e_i^\top + e_j e_j^\top$ and so,

$$\left\| \mathbb{E} \sum_{i,j \in [n] \text{ distinct}} (B^{(ij)})^2 \right\|_{\mathrm{op}} = \|(n-1)\mathrm{Id}\|_{\mathrm{op}} \leq n.$$

The matrix Bernstein inequality implies that

$$\mathbb{E}\,\|B\|_{\mathrm{op}} \leq O\left(\sqrt{\left\| \mathbb{E} \sum_{i,j \in [n] \text{ distinct}} (B^{(ij)})^2 \right\|_{\mathrm{op}}} \cdot \sqrt{\log n} + 2\log n \right) = O(\sqrt{n \log n}).$$

Markov's inequality implies that with high probability (say 0.99), $\|B\|_{\mathrm{op}} \leq O(\sqrt{n \log n})$, so

$$\left\| A - \frac{1}{2} J \right\|_{\mathrm{op}} = \leq \|B\|_{\mathrm{op}} + \left\| \frac{1}{2} \mathrm{Id} \right\|_{\mathrm{op}} = O(\sqrt{n \log n}),$$

completing the proof.

Problem 2 (Robustness to adversarial modification)**.** Suppose a malicious adversary is allowed to modify any subset of $n^{0.99}$ edges of a graph drawn from $G(n, 1/2)$. Show that, despite this, there exists with high probability a constant-degree SoS proof that certifies the graph does not contain any clique of size greater than $O(\sqrt{n \log n})$.

Solution We shall use essentially the same proof as in the first question. Let G be the true graph drawn from $G(n, 1/2)$ with adjacency matrix A, and $\widetilde{G}$ the corrupted graph observed by the algorithm $\widetilde{A}$. Then, we have

$$\left\|\widetilde{A} - \frac{1}{2}J\right\|_{\text{op}} \leq \left\|A - \frac{1}{2}J\right\|_{\text{op}} + \left\|\widetilde{A} - A\right\|_{\text{op}}.$$

However, because the adversary can modify only $n^{0.99}$ edges (that is, change $n^{0.99}$ entries of A from 0 to 1 or vice-versa),

$$\left\|\widetilde{A} - A\right\|_{\text{op}} \leq \left\|\widetilde{A} - A\right\|_{F} \leq n^{0.99/2} = o(\sqrt{n}).$$

Thus, with high probability, $\left\|A - \frac{1}{2}J\right\|_{\text{op}} \leq O(\sqrt{n \log n})$, and by the above argument, $\left\|\widetilde{A} - \frac{1}{2}J\right\|_{\text{op}} \leq O(\sqrt{n \log n})$, completing the proof by the same argument as in the first question.

Problem 3 (Planted 2-XOR)**.** Let ϕ be a random instance of 2-XOR over $\{\pm 1\}$, sampled in the following way. First, choose $x^* \in \{\pm 1\}^n$. Then, for each $(i, j) \in [n]^2$, with probability $\frac{Cn \log n}{n^2}$,

1. with probability 0.99, add the constraint $x_i x_j = x_i^* x_j^*$ to ϕ, and
2. otherwise, add the constraint $x_i x_j = -x_i^* x_j^*$ to ϕ.

The resulting instance ϕ should have about $Cn \log n$ equations.

(a) Show that for sufficiently large (constant) C, with high probability, there exists $y \in \{\pm 1\}^n$ which satisfies a 0.98 fraction of the equations in ϕ.

(b) Show that for sufficiently large (constant) C, there is a polynomial-time algorithm which finds some $y \in \{\pm 1\}^n$ which satisfies at least a 0.97 fraction of the equations in ϕ.

Solution

(a) We shall show that with high probability, x^* satisfies a 0.98 fraction of the equations in ϕ. First off, we may use the Chernoff bound to show that with high probability, there are about $Cn \log n(1 - o(1))$ clauses. Indeed, the number of clauses is distributed as the binomial random variable $\mathrm{Bin}(n^2, \frac{Cn \log n}{n^2})$. Then,

$$\Pr\left[|\#\text{ clauses} - Cn \log n| \geq n\sqrt{\log n}\right] \leq \exp\left(-O\left(\frac{(n\sqrt{\log n})^2}{n^2}\right)\right) = o(1).$$

Now, suppose we have conditioned on there being $m = (Cn \log n)(1 - o(1))$ clauses. Then, the number of clauses satisfied by x^* is distributed as $\mathrm{Bin}(m, 0.99)$. Again, a Chernoff bound implies that

$$\Pr\left[\#\text{ clauses satisfied by } x^* \leq 0.98m\right] \leq \exp\left(-O\left(\frac{(0.01m)^2}{m}\right)\right) = o(1).$$

Putting these two together, we get that

$$\Pr\left[\#\text{ clauses} = Cn \log n(1 + o(1)) \text{ and fraction of clauses satisfied by } x^* \geq 0.98\right] = o(1),$$

completing the proof. Note that the 0.98 here can be replaced with a constant arbitrarily close to 0.99. We shall use this in the second part.

(b) Suppose that the set of constraints is $\{x_i x_j = A_{ij} : ij \in E\}$, where each $A_{ij} \in \{\pm 1\}$ and E is the set of pairs involved in constraints. Let A be the corresponding constraint matrix, whose ijth entry is A_{ij} if $ij \in E$, and is 0 otherwise.This 2-XOR instance may naturally be represented as an optimization problem, where the goal is to maximize $\frac{1}{m} x^\top A x$ subject to the constraints $x_i^2 = 1$ for all i.

Consider the natural degree-2 sum-of-squares relaxation of the above optimization problem, and suppose it returns a pseudoexpectation $\widetilde{\mathbb{E}}$ over random variables x_i, with $\widetilde{\mathbb{E}} \vDash x_i^2 = 1$ for all i, and $\widetilde{\mathbb{E}} \frac{1}{m} x^\top A x \geq (0.98 - \varepsilon)$. We know that such a pseudoexpectation exists with high probability by the strengthened version of (a)—note that we have a $0.98 - \varepsilon$ instead of a $0.99 - \varepsilon$ here because $x^\top A x$ is equal to (#satisfied clauses) − (#unsatisfied clauses).

The algorithm will be to take the top eigenvector of $\widetilde{\mathbb{E}}\, xx^\top$, and round it to a vector on the hypercube by taking the sign of each coordinate of this vector.

We shall complete the proof in two steps. First, we will show that the top eigenvector y of $\widetilde{\mathbb{E}}\, xx^\top$ is well-correlated with x^*, and has large quadratic form $y^\top Ay$. Next, we will show that rounding y to a vector on the hypercube does not lose too much in the objective value.

Let us begin with the second of these steps: let y be a vector of norm $\sqrt{n}$ such that $\langle y, x^*\rangle \geq (1-\varepsilon)n$. Then,

$$\|y - x^*\|^2 \leq 2\varepsilon n,$$

so there are at most $2\varepsilon n$ indices i such that $|y_i - x_i^*| \geq 1$. In particular, this implies that there are most $2\varepsilon n$ indices i such that $\mathrm{sign}(y_i) \neq \mathrm{sign}(x_i^*)$, and thus $\langle \mathrm{sign}(y), x^*\rangle \geq (1-2\varepsilon)n$, completing the proof.

To conclude, we must show that the top eigenvector of $\widetilde{\mathbb{E}}\, xx^\top$ is well-correlated with x^*. We may write

$$(0.98-\varepsilon)m \leq \widetilde{\mathbb{E}}\, x^\top Ax = 0.98 \cdot \frac{C\log n}{n}\, \widetilde{\mathbb{E}}\langle x, x^*\rangle^2 + \left\langle \widetilde{\mathbb{E}}\, xx^\top, A - 0.98 \cdot \frac{C\log n}{n}(x^*)(x^*)^\top \right\rangle. \quad (2)$$

Indeed, observe that $\mathbb{E}[A \mid x^*] = 0.98 \cdot \frac{C\log n}{n} \cdot (x^*)(x^*)^\top$. Next, we shall show that with high probability, $\left\|A - 0.98 \cdot \frac{C\log n}{n} \cdot (x^*)(x^*)^\top\right\|_{\mathrm{op}}$ is small.

This operator norm bound essentially follows from the matrix Bernstein inequality. Indeed, let B be this matrix, and set $B^{(ij)}$ to be the symmetric random matrix such that

$$B^{(ij)}_{ij} = B^{(ij)}_{ji} = \begin{cases} -0.98 \cdot \frac{C\log n}{n} \cdot x_i^* x_j^*, & \text{w.p. } 1 - \frac{C\log n}{n}, \\ \left(1 - 0.98\frac{C\log n}{n}\right) x_i^* x_j^*, & \text{w.p. } 0.99 \cdot \frac{C\log n}{n}, \\ -\left(1 + 0.98\frac{C\log n}{n}\right) x_i^* x_j^*, & \text{w.p. } 0.01 \cdot \frac{C\log n}{n} \end{cases}$$

and all other entries are equal to 0. Clearly, the operator norm of $B^{(ij)}$ is almost surely bounded by 2, and

$$\begin{aligned} \mathbb{E}(B^{(ij)})^2 = (e_{ii} + e_{jj}) \cdot &\left(1 - \frac{C\log n}{n}\right) \cdot \left(0.98\frac{C\log n}{n}\right)^2 + \left(0.99 \cdot \frac{C\log n}{n}\right) \cdot \left(1 - 0.98\frac{C\log n}{n}\right)^2 \\ &+ \left(0.01 \cdot \frac{C\log n}{n}\right) \cdot \left(1 + 0.98\frac{C\log n}{n}\right)^2 \\ \leq (e_{ii} + e_{jj})&\frac{C\log n}{n}\left(1 + o(1)\right). \end{aligned}$$

It follows that

$$\left\|\sum_{i,j} (B^{(ij)})^2\right\|_{\mathrm{op}} \leq O(C\log n).$$

The matrix Bernstein inequality implies that

$$\mathbb{E}\left\|A - 0.98 \cdot \frac{C\log n}{n}(x^*)(x^*)^\top\right\|_{\mathrm{op}} \leq O\left(\sqrt{C}\log n\right).$$

For a sufficiently large constant C, this implies that with high probability, $\left\|A - 0.99 \cdot \frac{C\log n}{n}(x^*)(x^*)^\top\right\|_{\mathrm{op}} \leq 0.001C\log n$, and thus, since $m = Cn\log n(1+o(1))$, we can plug this back into (2) to conclude that

$$\widetilde{\mathbb{E}}\langle x, x^*\rangle^2 \geq (1-\varepsilon)n^2.$$

for a constant ε that can be made arbitrarily small. To complete the proof, we must use the above to show that the top eigenvector of $\widetilde{\mathbb{E}}\, xx^\top$ is correlated with x^*. Indeed, the above implies that

$$\begin{aligned}
\left\|\widetilde{\mathbb{E}}\, xx^\top - (x^*)(x^*)^\top\right\|_{\mathrm{op}} &\leq \left\|\widetilde{\mathbb{E}}\, xx^\top - (x^*)(x^*)^\top\right\|_F^2 \\
&\leq \left\|\widetilde{\mathbb{E}}\, xx^\top\right\|_F^2 + n^2 - \widetilde{\mathbb{E}}\langle x, x^*\rangle^2 \\
&\leq 2n^2 - \widetilde{\mathbb{E}}\langle x, x^*\rangle^2 \leq 2\varepsilon n^2.
\end{aligned}$$

Standard inequalities discussed in class (the Davis-Kahan Theorem, for instance) show that this implies that the top eigenvector of $\widetilde{\mathbb{E}}\, xx^\top$ is well-correlated with x^*, completing the proof.