Problem Set 3 Solutions

Sum of Squares Seminar

October 24, 2017

Problem 1: SOS Proofs (20 points)

Give sum of squares proofs for the following facts (over $\mathbb{R}$):

(a) 5 points: $\forall x, y, z, w, 4xyzw \leq x^4 + y^4 + z^4 + w^4$ (this is essentially the AM-GM inequality on 4 terms)

(b) 5 points: If $\sum_{i=1}^{n} x_i^2 = 1$ then $\sum_{i=1}^{n} x_i \leq \sqrt{n}$

(c) 10 points: If $\sum_{i=1}^{n} x_i^2 = 1$ then $\prod_{i=1}^{n} x_i^2 \leq n^{-n}$ (hint is available)

(d) Challenge question: If M is a doubly stochastic $n \times n$ matrix (i.e. all entries are nonnegative, all rows sum to one, and all columns sum to 1), then the permanent of M is at least $\frac{n!}{n^n}$

Solutions

(a) $x^4 + y^4 + z^4 + w^4 = (x^2 - y^2)^2 + (z^2 - w^2)^2 + 2(xy - zw)^2 + 4xyzw \geq 4xyzw$

(b) $\sum_{i=1}^{n} x_i = \sqrt{n} - \frac{1}{2}\sum_{i=1}^{n} (\sqrt[4]{n}x_i - \frac{1}{\sqrt[4]{n}})^2 - \frac{\sqrt{n}}{2}(1 - \sum_{i=1}^{n} x_i^2) \leq \sqrt{n}$

(c) We show by induction that $\forall k \in [1, n]$, $\frac{1}{\binom{n}{k}} \sum_{i_1 < \cdots < i_k} \prod_{j=1}^{k} x_{i_j}^2 \leq \frac{1}{n^k}$. The base case $k = 1$ is trivial. Now observe that since $\sum_{i=1}^{n} x_i^2 = 1$,

$$\begin{aligned}
\sum_{i_1 < \cdots < i_k} \prod_{j=1}^{k} x_{i_j}^2 &= (\sum_{i=1}^{n} x_i^2) \sum_{i_1 < \cdots < i_k} \prod_{j=1}^{k} x_{i_j}^2 \\
&= (k+1) \sum_{i_1 < \cdots < i_{k+1}} \prod_{j=1}^{k+1} x_{i_j}^2 + \sum_{i} \sum_{i_1 < \cdots < i_{k-1} : i_1, \cdots, i_{k-1} \neq i} x_i^4 \prod_{j=1}^{k-1} x_{i_j}^2 \\
&= (k+1) \sum_{i_1 < \cdots < i_{k+1}} \prod_{j=1}^{k+1} x_{i_j}^2 + \sum_{i_1 < \cdots < i_{k-1}} \sum_{i : i \notin \{i_1, \cdots, i_{k-1}\}} x_i^4 \prod_{j=1}^{k-1} x_{i_j}^2
\end{aligned}$$

Since $x_i^4 + x_{i'}^4 \geq 2x_i^2 x_{i'}^2$,

$$\sum_{i_1<\cdots<i_{k-1}} \sum_{i:i\notin\{i_1,\cdots,i_{k-1}\}} x_i^4 \prod_{j=1}^{k-1} x_{i_j}^2 \geq \frac{2}{n-k} \sum_{i_1<\cdots<i_{k-1}} \sum_{i,i':i<i',i,i'\notin\{i_1,\cdots,i_{k-1}\}} x_i^2 x_{i'}^2 \prod_{j=1}^{k-1} x_{i_j}^2$$

$$= \frac{2\binom{k+1}{2}}{n-k} \sum_{i_1<\cdots<i_{k+1}} \prod_{j=1}^{k+1} x_{i_j}^2$$

Putting these equations together,

$$\sum_{i_1<\cdots<i_k} \prod_{j=1}^{k} x_{i_j}^2 \geq \frac{(k+1)n}{n-k} \sum_{i_1<\cdots<i_{k+1}} \prod_{j=1}^{k+1} x_{i_j}^2$$

Thus,

$$\binom{n}{k+1} \sum_{i_1<\cdots<i_{k+1}} \prod_{j=1}^{k+1} x_{i_j}^2 \leq \frac{1}{n}\binom{n}{k} \sum_{i_1<\cdots<i_k} \prod_{j=1}^{k} x_{i_j}^2 \leq n^{-(k+1)}$$

where the last inequality follows from the inductive hypothesis.

Problem 2: Decomposing an L^1 pseudo-metric space (15 points)

Recall the objective function for the relaxation of sparsest cut:

$$\frac{\sum_{i<j,(i,j)\in E(G)} d_{ij}}{\sum_{i<j} d_{ij}}$$

Let G be the cycle on 6 vertices, i.e. $V(G) = v_1, \cdots, v_6$ and $E(G) = \{(v_i, v_{i+1}) : i \in [1,5]\} \cup \{(v_1, v_6)\}$. Assume that we are given the following mapping of $v_1, \cdots, v_6$ into $\mathbb{R}^2$:

$$v_1 = (0,1), v_2 = (1,0), v_3 = (2,0), v_4 = (3,0), v_5 = (3,2), v_6 = (0,2)$$

(a) 5 points: What is the value of the objective function given by this L^1 metric? What is the actual sparsity of G?

(b) 10 points: Decompose this L^1 metric as a positive linear combination of cut spaces. Which cut space(s) give the best value for the objective function?

Solutions

(a) The distances for this L^1 metric are as follows: $d(v_1, v_2) = 2$, $d(v_1, v_3) = 3$, $d(v_1, v_4) = 4$, $d(v_1, v_5) = 4$, $d(v_1, v_6) = 1$, $d(v_2, v_3) = 1$, $d(v_2, v_4) = 2$, $d(v_2, v_5) = 4$, $d(v_2, v_6) = 3$, $d(v_3, v_4) = 1$, $d(v_3, v_5) = 3$, $d(v_3, v_6) = 4$, $d(v_4, v_5) = 2$, $d(v_4, v_6) = 5$, and $d(v_5, v_6) = 3$.

Thus, $\sum_{i,j:i<j,(i,j)\ inE(G)} d(v_i, v_j) = 2 + 1 + 1 + 2 + 3 + 1 = 10$ and $\sum_{i,j:i<j} d(v_i, v_j) = 42$, giving an objective value of $\frac{10}{42} = \frac{5}{21}$

(b) Looking at the first coordinate, we have the following cuts, each with wieght 1

(a) $S = \{1, 6\}, \bar{S} = \{2, 3, 4, 5\}$

(b) $S = \{1, 2, 6\}, \bar{S} = \{3, 4, 5\}$

(c) $S = \{1, 2, 3, 6\}, \bar{S} = \{4, 5\}$

Looking at the second coordinate, we have the following cuts, each with wieght 1

(a) $S = \{2, 3, 4\}, \bar{S} = \{1, 5, 6\}$

(b) $S = \{1, 2, 3, 4\}, \bar{S} = \{5, 6\}$

The cuts $S = \{1, 2, 6\}, \bar{S} = \{3, 4, 5\}$ and $S = \{2, 3, 4\}, \bar{S} = \{1, 5, 6\}$ have the optimal sparsity.

Problem 3: Degree 4 Motzkin Polynomial Analgoue (15 points)

Consider the polynomial $p(x, y, z) = x^2y^2 + x^2z^2 + y^2z^2 - 4xyz + 1$.

(a) 5 points: Prove that $\forall x, y, z, p(x, y, z) \geq 0$

(b) 10 points: Prove that $p(x, y, z)$ cannot be written as the sum of squares of polynomials

Solutions

(a) Applying the AM-GM inequality on x^2y^2, x^2z^2, y^2z^2, and 1 we obtain that

$$\frac{1}{4}\left(x^2y^2 + x^2z^2 + y^2z^2 + 1\right) \geq \sqrt[4]{(x^2y^2)(x^2z^2)(y^2z^2)(1)} = xyz$$

Multiplying by 4 and rearranging, we obtain that $x^2y^2 + x^2z^2 + y^2z^2 - 4xyz + 1 \geq 0$, as needed.

(b) The Newton polytope P of $p(x, y, z)$ is the convex hull of $(0, 0, 0), (2, 2, 0), (2, 0, 2), (0, 2, 2)$, so $\frac{1}{2}P$ is the convex hull of $(0, 0, 0), (1, 1, 0), (1, 0, 1), (0, 1, 1)$. The only integer points in $\frac{1}{2}P$ are $(0, 0, 0), (1, 1, 0), (1, 0, 1), (0, 1, 1)$ so if $p(x, y, z)$ were a sum of squares of polynomials it would have to be a sum of terms of the form $(axy + bxz + cyz + d)^2$. However, there is no way for any such term to give a nonzero coefficient for xyz. Thus, $p(x, y, z)$ cannot be written as a sum of squares of polynomials, as needed.

Problem 4: SOS Proof of Cauchy-Schwarz with Expected Values (25 points)

Recall the Cauchy-Schwarz inequality: $\left(\sum_{i=1}^{n} x_i y_i\right)^2 \leq \left(\sum_{i=1} x_i^2\right)\left(\sum_{i=1} y_i^2\right)$. In lecture, we saw an SOS proof of one analogous statement about pseudo-expectation values, namely

$$\tilde{E}\left[\left(\sum_{i=1}^{n} x_i y_i\right)^2\right] \leq \tilde{E}\left[\left(\sum_{i=1} x_i^2\right)\left(\sum_{i=1} y_i^2\right)\right]$$

In this problem, we prove that there is also a sum of squares proof of the following alternative analogue of Cauchy-Schwarz:

$$\tilde{E}\left[\sum_{i=1}^{n} x_i y_i\right]^2 \leq \tilde{E}\left[\sum_{i=1} x_i^2\right]\tilde{E}\left[\sum_{i=1} y_i^2\right]$$

(a) 10 points: Prove that for any pseudo-expectation values $\tilde{E}$, $(\tilde{E}[xy])^2 \leq \tilde{E}[x^2]\tilde{E}[y^2]$

(b) 5 points: Deduce that for all i, j,

$$2\tilde{E}[x_i y_i]\tilde{E}[x_j y_j] \leq \tilde{E}[x_i^2]\tilde{E}[y_j^2] + \tilde{E}[x_j^2]\tilde{E}[y_i^2]$$

(c) 10 points: Use this to prove that for any pseudo-expectation values $\tilde{E}$,

$$\tilde{E}\left[\sum_{i=1}^{n} x_i y_i\right]^2 \leq \tilde{E}\left[\sum_{i=1} x_i^2\right]\tilde{E}\left[\sum_{i=1} y_i^2\right]$$

(hints are available)

Solutions

(a) Consider the pseudo-expectation value of $\left(\sqrt{\tilde{E}[y^2]}x \pm \sqrt{\tilde{E}[x^2]}y\right)^2$.

$$\tilde{E}\left[\left(\sqrt{\tilde{E}[y^2]}x \pm \sqrt{\tilde{E}[x^2]}y\right)^2\right] = 2\tilde{E}[x^2]\tilde{E}[y^2] \pm 2\sqrt{\tilde{E}[x^2]\tilde{E}[y^2]}\tilde{E}[xy] \geq 0$$

Unless $\tilde{E}[x^2] = 0$ or $\tilde{E}[y^2] = 0$, this implies that $\mp\tilde{E}[xy] \leq \sqrt{\tilde{E}[x^2]\tilde{E}[y^2]}$ which implies that $(\tilde{E}[xy])^2 \leq \tilde{E}[x^2]\tilde{E}[y^2]$, as needed. If $\tilde{E}[x^2] = 0$ or $\tilde{E}[y^2] = 0$ then from the in-class exercise in lecture, $\tilde{E}[xy] = 0$

(b) Using part a, $2\tilde{E}[x_i y_i]\tilde{E}[x_j y_j] \leq 2\sqrt{\tilde{E}[x_i^2]\tilde{E}[y_i^2]\tilde{E}[x_j^2]\tilde{E}[y_i^2]}$. Now note that

$$\left(\sqrt{\tilde{E}[x_i^2]\tilde{E}[y_i^2]} - \sqrt{\tilde{E}[x_j^2]\tilde{E}[y_i^2]}\right)^2 = \tilde{E}[x_i^2]\tilde{E}[y_i^2] + \tilde{E}[x_j^2]\tilde{E}[y_j^2] - 2\sqrt{\tilde{E}[x_i^2]\tilde{E}[y_i^2]\tilde{E}[x_j^2]\tilde{E}[y_i^2]} \geq 0$$

Combining these equations, $2\tilde{E}[x_i y_i]\tilde{E}[x_j y_j] \leq \tilde{E}[x_i^2]\tilde{E}[y_j^2] + \tilde{E}[x_j^2]\tilde{E}[y_i^2]$, as needed.

(c) Summing the equation in part b over all $i < j$,

$$\sum_{i,j:i<j} 2\tilde{E}[x_i y_i]\tilde{E}[x_j y_j] \leq \sum_{i,j:i<j} \left(\tilde{E}[x_i^2]\tilde{E}[y_j^2] + \tilde{E}[x_j^2]\tilde{E}[y_i^2]\right)$$

Applying the equation in part a with $x = x_i$ and $y = y_i$ and summing over all i,

$$\sum_i (\tilde{E}[x_i y_i])^2 \leq \sum_i \tilde{E}[x_i^2]\tilde{E}[y_i^2]$$

Adding these equations together, we obtain that

$$\tilde{E}\left[\sum_{i=1}^{n} x_i y_i\right]^2 = \left(\sum_i \tilde{E}[x_i y_i]\right)^2 \leq \left(\sum_i \tilde{E}[x_i^2]\right)\left(\sum_i \tilde{E}[y_i^2]\right) = \tilde{E}\left[\sum_{i=1} x_i^2\right]\tilde{E}\left[\sum_{i=1} y_i^2\right]$$

Problem 5: Reasoning Using Rational Functions (25 points)

Consider the constraint $(x^2+1)y = z^2$. We can immediately see that $y \geq 0$ as $y = \frac{z^2}{x^2+1}$ and both the numerator and the denominator must be non-negative. In this question, we consider whether the sum of squares hierarchy can capture this reasoning.

(a) 5 points: Give a sum of squares proof that if we add the constraint $y \leq -c$ (equivalently $y = -c - u^2$) for any $c > 0$ then the constraints are infeasible over $\mathbb{R}$.

(b) 20 points: Show that there exist degree 4 pseudo-expectation values $\tilde{E}$ with $\tilde{E}[y] < 0$ which respect the constraint that $(x^2+1)y = z^2$ (hint is available).

Solutions

(a) With the constraint $y = -c - u^2$, we have that

$$-c = (-c - u^2 - y)(x^2+1) + \big((x^2+1)y - z^2\big) + cx^2 + u^2x^2 + z^2$$

so these constraints are infeasible over $\mathbb{R}$

(b) It's easiest to find the pseudo-expectation values with a semidefinite program. To find such pseudo-expectation values by hand, we can do the following. We can use an actual distribution over solutions as a guide and then tweak it to get the desired pseudo-expectation values.

We take our distribution to have the following components. We choose the signs of x and z randomly with probability $\frac{1}{2}$, so any monomial which has odd degree in x or z automatically has expected value 0.

(a) For a small $\epsilon_1 > 0$ which will be chosen later, we take $|x| = \frac{1}{\epsilon_1}, y = 0, |z| = 0$ with probability ${\epsilon_1}^{3.9}$. This contributes ${\epsilon_1}^{-.1}$ to $E[x^4]$ and negligible amounts to everything else.

(b) For a small $\epsilon_2 > 0$ which will be chosen later, we take $|x| = \frac{1}{\epsilon_2}, y = \frac{1}{1+\epsilon_2^2}, |z| = \frac{1}{\epsilon_2}$ with probability ${\epsilon_2}^{3.9}$. This contributes ${\epsilon_2}^{-.1}$ to $E[x^4]$, $E[x^2z^2]$, and $E[z^4]$ and negligible amounts to everything else.

(c) For a small $\epsilon_3 > 0$ which will be chosen later, we take $|x| = \sqrt{\frac{1}{\epsilon_3} - 1}, y = \frac{1}{\epsilon_3}, |z| = \frac{1}{\epsilon_3}$ with probability ${\epsilon_3}^{3.9}$. This contributes ${\epsilon_3}^{-.1}$ to $E[y^4]$, $E[y^2z^2]$, and $E[z^4]$ and negligible amounts to everything else.

(d) For a small $\epsilon'_1 > 0$ which will be chosen later, we take $|x| = \frac{1}{\epsilon'_1}, y = 0, |z| = 0$ with probability ${\epsilon'_1}^{1.9}$. This contributes ${\epsilon'_1}^{-.1}$ to $E[x^2]$ and negligible amounts to all other terms of degree at most 2.

(e) For a small $\epsilon'_3 > 0$ which will be chosen later, we take $|x| = \sqrt{\frac{1}{\epsilon'_3} - 1}, y = \frac{1}{\epsilon'_3}, |z| = \frac{1}{\epsilon'_3}$ with probability ${\epsilon'_3}^{1.9}$. This contributes ${\epsilon'_3}^{-.1}$ to $E[y^2]$ and $E[z^2]$ and negligible amounts to all other terms of degree at most 2.

We now take $\epsilon_1 << \epsilon_2 << \epsilon_3 << {\epsilon'_1}^3 = {\epsilon'_3}^3$. We have that $E[y] = 0$ and the resulting moment matrix M is very much PSD, so much so that if we subtract 1 from the entries corresponding to y and z^2, M remains PSD. This now gives us $\tilde{E}[y] = -1$, as needed.

Hints

1c. More generally, show that $\forall k \in [1, n]$, $\frac{1}{\binom{n}{k}} \sum_{i_1 < \cdots < i_k} \prod_{j=1}^{k} x_{i_j}^2 \leq \frac{1}{n^k}$ by showing that if the inequality holds for k_1 and k_2 then it holds for $k_1 + k_2$ whenever $k_1 + k_2 \leq n$

4a. Consider the pseudo-expectation value of a square whose coefficients are functions of $\tilde{E}[x^2]$, $\tilde{E}[y^2]$, and/or $\tilde{E}[xy]$.

4b. Use part a to show that $2\tilde{E}[x_i y_i]\tilde{E}[x_j y_j] \leq 2\sqrt{\tilde{E}[x_i^2]\tilde{E}[y_i^2]\tilde{E}[x_j^2]\tilde{E}[y_i^2]}$

5. One way to find such pseudo-expectation values is to start with an actual expectation over distribution of solutions and then show that you can change the value of $\tilde{E}[y]$ to be negative. For this, it is useful to choose the distribution to make the diagonal entries very large without making the rest of the matrix too large. For example, if your distribution sets $x = B$, $y = z = 0$ with probability $B^{-3.5}$ where B is a large constant, this contributes $B^{.5}$ to $E[x^4]$ and almost nothing to the expected value of any other degree 4 monomial. Thus, we can make the x^4 entry arbitrarily large with negligible effect on the rest of the matrix.

Alternatively, you can write a semidefinite program to find such pseudo-expectation values.