Assignment 1

CPSC 663: Sum-of-Squares Optimization (Spring 2022)

Assigned: February 12, 2022   Due: March 4, 2022

Solve any four out of the five problems. If you solve them all, I will grade the first four. Each problem will be worth an equal amount towards your grade.

Problem 1 (Non-negative principal component analysis)**.** Suppose $\boldsymbol{S} \in \mathbb{R}^{n\times n}$ is symmetric. Consider the optimization problem

$$\begin{array}{ll} \text{maximize} & \boldsymbol{x}^\top \boldsymbol{S}\boldsymbol{x} \\ \text{subject to} & \sum_{i=1}^n x_i^2 = 1, \\ & x_i \geq 0 \text{ for all } i \in [n]. \end{array} \tag{1}$$

This kind of problem is sometimes called *positive* or *non-negative PCA*.

1. Formulate the degree 2 Lasserre relaxation, as an explicit matrix SDP, with constraints based on Putinar's Positivstellensatz.
2. Your SDP from Part 1 should have an $(n+1)\times(n+1)$ decision variable. Show that this dimension may be reduced to $n \times n$ without changing the value of the relaxation.
3. Explicitly give an optimal point of this reduced relaxation, as a function of $\boldsymbol{S}$.
4. Now, repeat Parts 1 and 2 for the degree 2 Lasserre relaxation with constraints based on Schmüdgen's Positivstellensatz.
5. Give an example of $\boldsymbol{S}$ for which the upper bound given by the SDP from Part 4 is strictly smaller than that given by the SDP from Part 2.

Problem 2 (One-dimensional optimization)**.** Let $p, q \in \mathbb{R}[x]$ be univariate polynomials with $\deg p, \deg q \leq d$.

1. Describe an algorithm that uses SDP to compute the global minimum of $p(x)$ based on a Parrilo relaxation, that runs in polynomial time in d. Write the program your algorithm solves as an explicit SDP with coefficients in terms of the coefficients of p.
2. Adjust your algorithm to do the same for the univariate rational function $p(x)/q(x)$, supposing $q(x)$ has no real zeroes.

Problem 3 (Nesterov's approximation algorithm)**.** In class we have discussed *Gaussian rounding* of SOS relaxations. Suppose we have a pseudoexpectation $\widetilde{\mathbb{E}}$, and we assume for the sake of simplicity that $\widetilde{\mathbb{E}}[\boldsymbol{x}] = 0$ for this problem. In Gaussian rounding, we then produce a point $\boldsymbol{y} \in \mathbb{R}^n$ by sampling $\boldsymbol{y} \sim \mathcal{N}(\boldsymbol{0}, \widetilde{\mathbb{E}}[\boldsymbol{x}\boldsymbol{x}^\top])$. In this problem we consider what happens if we want instead to produce a point $\boldsymbol{y} \in \{\pm 1\}^n$. The natural adjustment of Gaussian rounding is to take

$$\boldsymbol{g} \sim \mathcal{N}(\boldsymbol{0}, \widetilde{\mathbb{E}}[\boldsymbol{x}\boldsymbol{x}^\top]), \tag{2}$$

$$y_i := \mathsf{sgn}(g_i) \text{ for each } i \in [n]. \tag{3}$$

1. Show that, for each $i, j \in [n]$,

$$\mathbb{E}[y_i y_j] = \frac{2}{\pi} \sin^{-1}\left(\widetilde{\mathbb{E}}[x_i x_j]\right) \tag{4}$$

2. Prove the matrix inequality

$$\mathbb{E}[\boldsymbol{y}\boldsymbol{y}^\top] \succeq \frac{2}{\pi}\widetilde{\mathbb{E}}[\boldsymbol{x}\boldsymbol{x}^\top]. \tag{5}$$

   Use the Schur product theorem: if $\boldsymbol{S}, \boldsymbol{T} \succeq \boldsymbol{0}$, then the matrix $\boldsymbol{M}$ with $M_{ij} = S_{ij}T_{ij}$ also has $\boldsymbol{M} \succeq \boldsymbol{0}$.

3. Consider the *little Grothendieck problem*: given $\boldsymbol{A} \succeq \boldsymbol{0}$, we want to solve the optimization

$$\begin{array}{ll} \text{maximize} & \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} \\ \text{subject to} & \boldsymbol{x} \in \{\pm 1\}^n. \end{array} \tag{6}$$

   Use the previous parts to describe a randomized polynomial-time $\frac{2}{\pi}$-approximation algorithm for this problem.[1] Why doesn't your algorithm give the same approximation for arbitrary $\boldsymbol{A}$?

Problem 4 (SOS proofs)**.** Give SOS proofs of the following inequalities.

1. (AM-GM inequality) Prove that, for any $n \geq 1$,

$$\frac{x_1^{2n} + \cdots + x_n^{2n}}{n} - x_1^2 \cdots x_n^2 \in \mathsf{SOS}. \tag{7}$$

   To do this, let S_n denote the permutations of $[n]$, and write $\sigma(i)$ for $i \in [n]$ and $\sigma \in S_n$ as the image of i under σ. Define $p_k(\boldsymbol{x}) := \frac{1}{n!}\sum_{\sigma \in S_n} x_{\sigma(1)}^{2(n-k)} x_{\sigma(2)}^2 \cdots x_{\sigma(2+k-1)}^2$ for $k = 0, \ldots, n-1$. Show that the left-hand side above is $p_0(\boldsymbol{x}) - p_{n-1}(\boldsymbol{x})$, and show that $p_{k-1}(\boldsymbol{x}) - p_k(\boldsymbol{x}) \in \mathsf{SOS}$ for each $k \in [n-1]$.

[1] Note that MaxCut is a special case of the little Grothendieck problem, since any graph Laplacian is psd. The Goemans-Williamson analysis for MaxCut uses special properties of graph Laplacians that we do not use in the analysis outlined here, and by taking advantage of those gets the stronger 0.878^+-approximation.

2. (Gaussian moments) Prove that, for any $n \geq 1$,

$$(2n)^n \|\boldsymbol{x}\|_2^{2n} - \mathop{\mathbb{E}}_{\boldsymbol{g} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})} \langle \boldsymbol{g}, \boldsymbol{x} \rangle^{2n} \in \mathsf{SOS}. \tag{8}$$

Note that the second term should be viewed as a polynomial in $\boldsymbol{x}$, and the expectation is over $\boldsymbol{g}$ a random Gaussian vector with i.i.d. standard Gaussian entries. You will have to compute the moments of these entries—use integration by parts and induction to prove that $\mathbb{E} g_i^{2k} = (2k-1)!! = 1 \cdot 3 \cdot 5 \cdots (2k-1)$.

3. (Triangle inequalities) Find $q_i(\boldsymbol{x})$ for $i \in [3]$ and $\boldsymbol{x} = (x_1, x_2, x_3)$ such that

$$x_1 x_2 + x_2 x_3 + x_1 x_3 + 1 \in q_1(\boldsymbol{x})(x_1^2 - 1) + q_2(\boldsymbol{x})(x_2^2 - 1) + q_3(\boldsymbol{x})(x_3^2 - 1) + \mathsf{SOS}. \tag{9}$$

Recall that these correspond to inequalities we included in the LP relaxation of MaxCut that we talked about, but which were omitted from the Goemans-Williamson relaxation. This problem shows that a sufficiently high degree SOS relaxation (degree 4 will suffice) does automatically satisfy these inequalities.

Problem 5 (Newton polytopes and Motzkin revisited)**.** Recall the Motzkin polynomial,

$$f(x, y) = x^4 y^2 + x^2 y^4 + 1 - 3x^2 y^2. \tag{10}$$

In class we gave a rather ad hoc proof that $m \notin \mathsf{SOS}$. In this exercise you will see a general framework that helps with many such proofs.

To a monomial in $(x_1, \dots, x_n)$, say $x_1^{a_1} \cdots x_n^{a_n}$, we associate the vector $(a_1, \dots, a_n) \in \mathbb{N}^n$. The *Newton polytope* of $p \in \mathbb{R}[x_1, \dots, x_n]$, denoted $N_p \subset \mathbb{R}^n$, is the convex hull of the vectors associated to the monomials having non-zero coefficients in p. Note that N_p only depends on which monomials appear in p, not on their coefficients.

1. Suppose $p(\boldsymbol{x}) = \sum_{i=1}^m q_i(\boldsymbol{x})^2$. Show that, for each i, N_{q_i} is contained in $\frac{1}{2} N_p$.

   Hint: Consider K the convex hull of all vectors associated to all monomials in all q_i, or equivalently the convex hull of all points in all of the N_{q_i}. Note that vertices of some of the N_{q_i} might not be vertices of K. Show that if $(a_1, \dots, a_n)$ is a vertex of K, then $x_1^{2a_1} \cdots x_n^{2a_n}$ must appear with a strictly positive coefficient in $p(\boldsymbol{x})$.

2. Draw N_f, the Newton polytope of the Motzkin polynomial, over the $\mathbb{N}^2$ grid. What monomials can appear in the $q_i(\boldsymbol{x})$ in a hypothetical SOS expression for f?

3. Using the restriction derived in Part 2, finish the proof that $f \notin \mathsf{SOS}$.

4. Suppose we are promised that $p(\boldsymbol{x}) \in \mathsf{SOS}$, and are interested in finding an explicit SOS decomposition using SDP. Sketch how you could improve the efficiency of this SDP using the properties of Newton polytopes. Roughly speaking, for high-degree polynomials, what geometric quantity associated to the Newton polytope governs the complexity of the improved SDP?