Assignment 2

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

Assigned: March 14, 2022   Due: April 4, 2022

Solve Problem 1, and any two out of the three remaining problems. If you solve them all, I will grade Problems 2 and 3. Each problem will be worth an equal amount towards your grade.

Problem 1 (Project topic)**.** Propose a paper to read or an open problem to discuss for your final project.

If you plan to read a paper, tell me the title and authors, read the abstract and introduction, and describe in one paragraph how it relates to SOS, what aspect of the paper you are interested in, and what else you might have to read to understand it.

If you plan to work on an open problem, write down the problem, give a few references concerning it, and outline in one paragraph what approach you plan to try or what numerical experiments you plan to run.

In either case, your final project will consist of a presentation on the last day of class (20-30 minutes) and a short write-up (1-2 pages) about whatever you choose here.

Problem 2 (Eigenvector perturbation bound)**.** Suppose $\boldsymbol{M} \succeq \boldsymbol{0}$, and $\boldsymbol{\Delta}$ has the same dimensions as $\boldsymbol{M}$ with $\|\boldsymbol{\Delta}\| < \lambda_1(\boldsymbol{M}) - \lambda_2(\boldsymbol{M})$ (the matrix norm without a subscript always denotes the operator norm). Let $\boldsymbol{v}$ be the top eigenvector of $\boldsymbol{M}$ and $\widetilde{\boldsymbol{v}}$ the top eigenvector of $\boldsymbol{M} + \boldsymbol{\Delta}$ (so that both are unit vectors). You will show the perturbation inequality

$$\langle \boldsymbol{v}, \widetilde{\boldsymbol{v}} \rangle^2 \geq 1 - \left(\frac{\|\boldsymbol{\Delta}\|}{\lambda_1(\boldsymbol{M}) - \lambda_2(\boldsymbol{M}) - \|\boldsymbol{\Delta}\|} \right)^2. \tag{1}$$

In our SOS arguments, we have been implicitly invoking something like this when rounding pseudoexpectations by taking the top eigenvector of the degree 2 pseudomoment matrix. Follow these steps:

1. Show that $\lambda_1(\boldsymbol{M}) - \lambda_i(\boldsymbol{M} + \boldsymbol{\Delta}) \geq \lambda_1(\boldsymbol{M}) - \lambda_2(\boldsymbol{M}) - \|\boldsymbol{\Delta}\|$ for all $i \geq 2$.
2. Using Step 1, show that $\|\boldsymbol{\Delta v}\| \geq (\lambda_1(\boldsymbol{M}) - \lambda_2(\boldsymbol{M}) - \|\boldsymbol{\Delta}\|) \cdot \|(\boldsymbol{I} - \widetilde{\boldsymbol{v}}\widetilde{\boldsymbol{v}}^\top)\boldsymbol{v}\|$.
3. Complete the proof.

Problem 3 (2 → 4 norm and small set expansion)**.** Let $G = (V, E)$ be a d-regular graph with adjacency matrix $\boldsymbol{A}$. Fix some $\alpha \in [0, 1]$, and let U be the subspace spanned by all eigenvectors of $\boldsymbol{A}$ with eigenvalue at least αd. Let $S \subseteq V$.

Suppose we draw $\boldsymbol{x} \sim \mathsf{Unif}(S)$ and then $\boldsymbol{y}$ a uniformly random neighbor of $\boldsymbol{x}$. We then define

$$\Phi(S) := \mathbb{P}[\boldsymbol{y} \notin S], \tag{2}$$

a measure of the *expansion* of S. Show that

$$\Phi(S) \geq 1 - \alpha - \|U\|_{2\to 4}^2 \sqrt{|S|}. \tag{3}$$

> HINT: First, show that $\|U\|_{2\to 4} = \|\boldsymbol{P}\|_{2\to 4} = \|\boldsymbol{P}\|_{4/3\to 2}$, where $\boldsymbol{P}$ is the projection matrix to U. Then, let $\boldsymbol{v}$ be the indicator vector of S, $v_i = \mathbb{1}\{i \in S\}$. Express $\Phi(S)$ in terms of a quadratic form with $\boldsymbol{v}$ and $\boldsymbol{A}$, and consider the decomposition of $\boldsymbol{v}$ into components in U and $U^\perp$.

This shows that when the 2 → 4 norm of the top eigenspace of a graph is small then the graph is a *small set expander*, i.e., sufficiently small subsets of vertices have large expansion. In fact, a converse is also true, which has been used to show that a good approximation of the 2 → 4 norm would refute the *Small Set Expansion Hypothesis*, which states, roughly speaking, that it is hard to distinguish good small set expanders from bad ones. This gives some evidence that it should be hard to approximate the 2 → 4 norm (and other similar quantities) in the worst case.

Problem 4 (Association schemes)**.** Recall the definition of an association scheme. We have matrices $\boldsymbol{I}_n = \boldsymbol{A}_1, \boldsymbol{A}_2, \dots, \boldsymbol{A}_m \in \{0, 1\}^{n\times n}_{\mathsf{sym}}$ where all $\boldsymbol{A}_i$ are non-zero and satisfy the following properties:

$$\sum_{j=1}^{m} \boldsymbol{A}_j = \boldsymbol{1}\boldsymbol{1}^\top, \tag{4}$$

$$\boldsymbol{A}_i \boldsymbol{A}_j = \sum_{k=1}^{m} c_{ijk} \boldsymbol{A}_k \text{ for some } c_{ijk} \in \mathbb{R}. \tag{5}$$

Recall that (5) implies that the $\boldsymbol{A}_i$ commute, and so are simultaneously diagonalizable. That is, there exist projection matrices $\boldsymbol{P}_1, \dots, \boldsymbol{P}_d$ to mutually orthogonal subspaces of $\mathbb{R}^n$ and some λ_{ij} such that

$$\boldsymbol{A}_i = \sum_{j=1}^{d} \lambda_{ij} \boldsymbol{P}_j, \tag{6}$$

so that $\lambda_{i1}, \dots, \lambda_{id}$ are the eigenvalues of $\boldsymbol{A}_i$. In this problem we will show how one can find the λ_{ij} from the c_{ijk}.

1. Show that the $\boldsymbol{A}_i$ are linearly independent.

2. Show that if $\boldsymbol{A}$ is a symmetric matrix and $\boldsymbol{P}$ is the orthogonal projection to the eigenspace of an eigenvalue λ, then $\boldsymbol{P}$ is a polynomial in $\boldsymbol{A}$. Conclude that we may take $d = m$ above (i.e., the number of distinct eigenspaces of each $\boldsymbol{A}_i$ is at most the total number of $\boldsymbol{A}_i$ in the scheme, as we saw in the Johnson scheme example in class), and that in this case the $\boldsymbol{P}_j$ are a basis for the span of the $\boldsymbol{A}_i$.

3. Show that $\lambda_{ik}\lambda_{jk} = \sum_{\ell=1}^{d} c_{ij\ell}\lambda_{\ell k}$.

4. Let $\boldsymbol{E} \in \mathbb{R}^{m\times m}$ have $E_{ij} = \lambda_{ji}$. Let $\boldsymbol{L}_i \in \mathbb{R}^{m\times m}$ have $(\boldsymbol{L}_i)_{kj} = c_{ijk}$. Show that $\boldsymbol{E}$ is non-singular, and that
$$\boldsymbol{E}\boldsymbol{L}_i\boldsymbol{E}^{-1} = \mathrm{diag}(\lambda_{i1}, \dots, \lambda_{im}). \tag{7}$$
That is, the distinct eigenvalues of $\boldsymbol{A}_i \in \mathbb{R}^{n\times n}_{\mathrm{sym}}$ are the eigenvalues of $\boldsymbol{L}_i \in \mathbb{R}^{m\times m}_{\mathrm{sym}}$, usually a much smaller matrix.