Sum of Squares seminar- Homework 0.

Here is some reading and exercises that I would like you to do before the course. Feel free to collaborate with others while solving those. You don't need to submit them, or even write the solutions down properly or anything — just make sure you know the material. Also, please don't hesitate to email me with any questions. (Please include "SOS course" in the email subject so I know that the email is related to the course.)

Background. I will assume general "mathematical maturity", and familiarity with some topics that are typically covered in undergraduate courses such as: eigenvectors and eigenvalues, linear programming, duality and Farkas lemma, basics of probability (expectation, variance, concentration), basic spectral graph theory (graphs and their adjacency matrices, relation between spectrum of adjacency matrix and random walk). Some sources for this material include Ryan O'Donnell's CMU class "15-859T: A Theorist's Toolkit" (available online on `http://www.cs.cmu.edu/~odonnell/toolkit13/`), see in particular Lectures 6–8 (spectral graph theory) and Lectures 13–14 (linear programming). See also the lecture notes for Jonathan Kelner's MIT course "18.409 Topics in Theoretical Comp Sci" (available online on `http://stellar.mit.edu/S/course/18/fa09/18.409/materials.html`). While not strictly necessary, you may find Luca Trevisan series of blog posts on expanders (from 2006, 2008, and 2011) illuminating, see `http://lucatrevisan.wordpress.com/tag/expanders/` .

Required reading. One thing I do want you to do before the seminar is to read my survey with David Steurer "Sum of Squares proofs and the quest toward optimal algorithms" at `http://eccc.hpi-web.de/report/2014/059/`. I believe it will be extremely helpful for you to follow the seminar.

Exercises

All matrices and vectors will be over the reals. In all the exercises below you can use the fact that any $n \times n$ matrix A has a singular value decomposition (SVD)

$$A = \sum_{i=1}^{r} \sigma_i u_i \otimes v_i$$

with $\sigma_i \in R$ and $u_i, v_i \in \mathbb{R}^n$, and for every i, j $\|u_i\| = 1$, $\|v_j\| = 1$ (where $\|v\| = \sqrt{\sum v_i^2}$), and for all $i \neq j$, $\langle u_i, u_j \rangle = 0$ and $\langle v_i, v_j \rangle = 0$. (For vectors u, v, their tensor product is defined as $u \otimes v$ is the matrix $T = uv^\top$ where $T_{i,j} = u_i v_j$.) Equivalently $A = U \Sigma V^\top$ where Σ is a diagonal matrix and U and V are orthogonal matrices (satisfying $U^\top U = V^\top V = I$). If A is symmetric then there is such a decomposition with $u_i = v_i$ for all i (i.e., $U = V$). In this case the values $\sigma_1, \ldots, \sigma_r$ are known as *eigenvalues* of A and the vectors $v_1, \ldots, v_r$ are known as *eigenvectors*. (This decomposition is unique if $r = n$ and all the σ_i's are distinct.) Moreover the SVD of A can be found in polynomial time. (You can ignore issues of numerical accuracy in all exercises.)

Exercise 1. For an $n \times n$ matrix A, the *spectral norm* of A is defined as the maximum of $\|Av\|$ over all vectors $v \in \mathbb{R}^n$ with $\|v\| = 1$.

1. Prove that if A is symmetric (i.e., $A = A^\top$), then $\|A\| \leq \max_i \sum_j |A_{i,j}|$. See footnote for hint[1]

2. Show that if A is the adjacency matrix of a d-regular graph then $\|A\| = d$.

Exercise 2. Let A be a symmetric $n \times n$ matrix. The *Frobenius norm* of A, denoted by $\|A\|_F$, is defined as $\sqrt{\sum_{i,j} A_{i,j}^2}$.

1. Prove that $\|A\| \leq \|A\|_F \leq \sqrt{n}\|A\|$. Give examples where each of those inequalities is tight.

2. Let $\mathrm{tr}(A) = \sum A_{i,i}$. Prove that for every even k, $\|A\| \leq \mathrm{tr}(A^k)^{1/k} \leq n^{1/k}\|A\|$.

3. (harder) Let A be a symmetric matrix such that $A_{i,i} = 0$ for all i and $A_{i,j}$ is chosen to be a random value in $\{\pm 1\}$ independently of all others. **(a)** Prove that (for n sufficiently large) with probability at least 0.99, $\|A\| \leq n^{0.9}$. **(b)** Prove that with probability at least 0.99, $\|A\| \leq n^{0.51}$.

Note: While $\|A\|$ can be computed in polynomial time, both $\max_i \sum_j |A_{i,j}|$ and $\|A\|_F$ give even simpler to compute upper bounds for $\|A\|$. However the examples in Exercise 1 and 2 show that they are not always tight. It is often easier to compute $\mathrm{tr}(A^k)^{1/k}$ than trying to compute $\|A\|$ directly, and as k grows this yields a better and better estimate.

Exercise 3. Let A be an $n \times n$ symmetric matrix. Prove that the following are equivalent:

1. A is positive semi-definite. That is, for every vector $v \in R^n$, $v^\top Av \geq 0$ (where we think of vectors as column vectors and so $v^\top Av = \sum_{i,j} A_{i,j}v_iv_j$).

2. All eigenvalues of A are non-negative. That is, if $Av = \lambda v$ then $\lambda \geq 0$.

3. The quadratic polynomial P_A defined as $P_A(x) = \sum A_{i,j}x_ix_j$ is a *sum of squares*. That is, there are linear functions $L_1, \ldots, L_m$ such that $P_A = \sum_i (L_i)^2$.

4. $A = B^\top B$ for some $n \times r$ matrix B

5. There exist a set of correlated random variables $(X_1, \ldots, X_m)$ such that for every i, j, $\mathbb{E}X_iX_j = A_{i,j}$ and moreover, for every i, the random variable X_i is distributed like a Normal variable with mean 0 and variance $A_{i,i}$.

Exercise 4. Give a polynomial-time algorithm that given a matrix A that is *not* positive semidefinite, finds a matrix M such that $\langle A, M\rangle < 0$, where $\langle A, M\rangle = \sum_{i,j} A_{i,j}M_{i,j} = \mathrm{tr}(AM)$ but $\langle B, M\rangle \geq 0$ for every B that is positive semidefinite. (Such an algorithm is known as a *separation oracle* for the set of positive semidefinite matrices.)

Exercise 5. Let d be even. Recall that a polynomial P of degree d is a *sum of squares* if there exist polynomials $Q_1, \ldots, Q_r$ such that $P = \sum Q_i^2$.

[1]**Hint:** You can do this via the following stronger inequality: for any (not necessarily symmetric) matrix A, $\|A\| \leq \sqrt{\alpha\beta}$ where $\alpha = \max_i \sum_j |A_{i,j}|$ and $\beta = \max_j \sum_i |A_{i,j}|$.

1. Prove that if P is a sum of squares, then in every such decomposition of it $\deg Q_i \leq d/2$ for all i. See footnote for hint[2]

2. We say that P is *homogenous* if every monomial of P has degree exactly d. Prove that if P is homogenous and a sum of squares then it has a decomposition where every Q_i is homogenous as well. See footnote for hint[3]

Exercise 6. For A an $n^2 \times n^2$ symmetric matrix, we let P_A be the degree 4 polynomial $P_A(x) = \sum_{i,j,k,\ell} A_{i,j,k,\ell} x_i x_j x_k x_l$. We say that $A \sim B$ if $P_A = P_B$.

1. Show that the set of B such that $B \sim A$ is an affine subspace of $\mathbb{R}^{n^2}$ (i.e., it is defined by linear equations on the coefficients).

2. Prove that P_A is a sum of squares polynomial if and only if there exists a positive semidefinite matrix B such that $B \sim A$.

3. (harder) Prove that P_A is a sum of squares polynomial if and only if there *does not* exist an $n^2 \times n^2$ matrix X such that for every permutation $\pi : [4] \to [4]$ and $i_1, \ldots, i_4 \in [n]$, $X_{i_1,i_2,i_3,i_4} = X_{i_{\pi(1)},i_{\pi(2)},i_{\pi(3)},i_{\pi(4)}}$, X is positive semidefinite, and $\mathrm{tr}(AX) < 0$. (This is semidefinite programming *duality*— can you see why?)

Exercise 7. Let $\mathcal{P}_d^n$ denote the set of n-variate polynomials of degree d and $\mathcal{S}_d^n$ denote the set of such polynomials that are sums of squares.

1. Prove that $\mathcal{P}_d^n$ is a linear subspace with dimension smaller than n^{2d}.

2. Prove that if $P, Q \in \mathcal{S}_d^n$ and $\alpha, \beta \geq 0$, then $\alpha P + \beta Q \in \mathcal{S}_d^n$.

3. Prove that if $P \in \mathcal{P}_d^n \setminus \mathcal{S}_d^n$ then there exists a linear function $L : \mathcal{P}_d^n \to \mathbb{R}$ such that $L(P) < 0$ but $L(Q) \geq 0$ for every $Q \in \mathcal{S}_d^n$. Can you give an $n^{O(d)}$ time algorithm to find such a function L given P? (This is a separation oracle for $\mathcal{S}_d^n$.)

Exercise 8. Prove that the following 4-variate polynomial is a sum of squares:

$$P(a,b,c,d) = \tfrac{1}{4}\left[a^8 + b^8 + c^8 + d^8\right] - a^2b^2c^2d^2$$

Exercise 9 (Harder - bonus)**.** Prove that the following 4-variate polynomial is *not* a sum of squares:

$$P(a,b,c,d) = \tfrac{1}{4}\left[a^2b^2 + a^2c^2 + b^2c^2 + d^4\right] - abcd$$

[2] **Hint:** Prove that the coefficient of the highest degree in the Q_i^2's is always positive and so can't be canceled.

[3] **Hint:** For every Q_i, let Q_i' denote the polynomial obtained by taking only the monomials of Q_i of degree $d/2$. Prove that $P = \sum Q_i'^2$.