Polynomial Optimization – Exercises

June 11, 2018

1. Let $G = ([n], E)$ be an undirected graph with vertex set $[n] = \{1, \ldots, n\}$ for a positive integer n and edge set E consisting of pairs of vertices. A set $S \subseteq [n]$ is said to be *stable* or *independent* if for any two vertices $i, j \in S$, the edge $ij \notin E$. Formulate a polynomial optimization problem to find the maximum cardinality stable set in G.

2. A *cut* in G is a partitioning of its vertices into two sets T and $[n] \backslash T$ and the size of the cut is the number of edges that go between the two parts. Formulate a polynomial optimization problem to find the maximum cardinality cut in G. This is another NP-hard problem.

3. A very common problem that arises in applications is to find the closest point in a given set from a given *data* point that has been observed in an experiment. For instance in computer vision one is often interested in reconstructing a three-dimensional scene from noisy images of the scene. The set of all true images that are possible by the given cameras is an algebraic set which is the *model* and the noisy images form the *data* point. If the noise model is Gaussian then the closest point to the model from the observed noisy data point is the maximum likelihood estimate. Model this problem as a polynomial optimization problem.

   Another problem that is very common in applications is to find a low rank estimate of a given matrix. Write down a polynomial optimization problem for finding the closest (in Euclidean distance) rank one real matrix of size $p \times q$ to a given real matrix A of the same size. Generalize to rank k. The classical *Eckart-Young theorem* in linear algebra gives a solution to this distance minimization problem. Look it up and see if you can solve it using the model you wrote.

4. A function f is convex if $f(\alpha x + \beta y) \leq \alpha f(x) + \beta f(y)$ for $x, y \in \mathbb{R}^n$ and scalars $\alpha, \beta \in \mathbb{R}$ such that $0 \leq \alpha, \beta$ and $\alpha + \beta = 1$. A set $K \subset \mathbb{R}^n$ is convex if for all $x, y \in K$ and $\alpha, \beta \in \mathbb{R}$ such that $0 \leq \alpha, \beta$ and $\alpha + \beta = 1$, any point of the form $\alpha x + \beta y \in K$. In other words, the line segment joining x and y is entirely in K.

   Consider the semialgebraic region $K = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$. Prove that K is a convex set if $-g_1, \ldots, -g_m$ are convex functions.

   If in addition f is a convex function, then the polynomial optimization problem $\min\{f(x) : x \in K\}$ is called a *convex program*.

5. (a) Convince yourself that the psd cone $\mathcal{S}^n_+ \subset \mathcal{S}^n$ is closed, convex, pointed and full-dimensional (solid). A cone with all these properties is called a *proper cone.*

   Recall that a *convex* cone $K \subset \mathbb{R}^t$ is one in which for every $x, y \in K$, $\lambda x + \mu y \in K$ for all $\lambda, \mu \geq 0$. The cone K is *pointed* if it does not contain any lines through the origin, i.e., there is no $x \in K$, $x \neq 0$ such that $-x \in K$.

   (b) Prove that the rank one matrices in $\mathcal{S}^n_+$ generate its *extreme rays* (i.e., rays that cannot be written as a non-negative combination of other rays in $\mathcal{S}^n_+$). Recall that a rank one matrix in $\mathcal{S}^n_+$ looks like $aa^\top$ where $a \in \mathbb{R}^n$.

   (c) By *Caratheodory's theorem* from convex geometry, every element in $\mathcal{S}^n_+$ can be written as a non-negative combination of at most $\frac{n(n+1)}{2}$ extreme rays of $\mathcal{S}^n_+$ On the other hand, the previous exercise allows you to bound the number of rank one matrices needed to write a psd matrix in $\mathcal{S}^n_+$ as a non-negative combination. How do these bounds compare?

6. Recall that the feasible region of a semidefinite program (SDP) is called a *spectrahedron.* We may take the following to be the official definition:

   Definition 0.1. A spectrahedron is a set of the form

   $$\{(x_1, \dots, x_m) \in \mathbb{R}^m \,:\, A_0 + \sum A_i x_i \geq 0\}$$

   where the matrices $A_i \in \mathcal{S}^n$.

   (a) In the lecture we defined a *spectrahedron* to be an affine slice of the psd cone. Indeed, the matrices defined by the above set is the intersection of the psd cone $\mathcal{S}^n_+$ with the affine plane obtained by translating $\operatorname{span}(A_1, \dots, A_m)$ by A_0. If the matrices $A_1, \dots, A_m$ are linearly independent in $\mathcal{S}^n$ then prove that there is a bijection between the two descriptions of a spectrahedron as a subset of $\mathbb{R}^m$ and $\mathcal{S}^n$ respectively.

   (b) Prove that a spectrahedron also admits the following descriptions:

      i. $\{X \in \mathcal{S}^n_+ \,:\, \langle B_j, X\rangle = b_i \;\forall\; j = 1, \dots, t\}$, for some symmetric matrices $B_j \in \mathcal{S}^n$,

      ii. $\{x \in \mathbb{R}^s \,:\, p_j(x) \geq 0 \;\; p_j \in \mathbb{R}[x_1, \dots, x_s],\; j = 1, \dots, r\}$

   How do t, s and r relate to m and n?

   (c) Using any of the above descriptions, argue that a spectrahedron is closed, convex and basic semi-algebraic.

   (d) Consider the following concrete spectrahedron:

   $$\mathcal{F} := \left\{ (x, y) \in \mathbb{R}^2 \,:\, \begin{bmatrix} x+1 & 0 & y \\ 0 & 2 & -x-1 \\ y & -x-1 & 2 \end{bmatrix} \geq 0 \right\}.$$

      i. Express $\mathcal{F}$ in the two other formats mentioned above.

ii. Draw this set in the plane.

iii. What is the polynomial that defines the boundary of $\mathcal{F}$? Generalize your result to the general spectrahedron in Definition 0.1.

7. A very common example of a spectrahedron is the *elliptope* $\mathcal{E}_n$ defined as follows.

$$\mathcal{E}_n := \{X \in \mathcal{S}_+^n : X_{ii} = 1 \ \forall\ i = 1, \dots, n\}.$$

   (a) What is the dimension of $\mathcal{E}_n$?

   (b) Use a computer to draw $\mathcal{E}_3$.

   (c) What are the rank one psd matrices on $\mathcal{E}_3$? Can you see them in your picture?

   (d) Find a rank two matrix on $\mathcal{E}_3$ that is not a convex combination of the rank one matrices on $\mathcal{E}_3$.

   (e) Can you model the max cut problem as an SDP over $\mathcal{E}_n$ with possibly additional rank constraints?

8. Check that the following basic facts are true for a sos polynomial $p = \sum h_j^2$ in $\mathbb{R}[x]$.

   (a) $\deg(p) = 2d \ \Rightarrow\ \deg(h_j) \le d$.

   (b) p homogeneous and $\deg(p) = 2d \ \Rightarrow\ h_j$ homogeneous and $\deg(h_j) = d$.

   (c) If $\tilde{p}$ is the homogenization of p then $p \ge 0$ (resp. sos) $\ \Leftrightarrow\ \tilde{p} \ge 0$ resp. sos.

   (d) If $\deg(p) = 2d$, bound the number of squares needed in the sos expression for p. (Hint: use that p sos if and only if $p = [x]_d^\top Q[x]_d$ for some $Q \succeq 0$.)

9. Write the following polynomial as a sos: $x^2 + 4x + 5$.

10. Express $2x^4 + 5y^4 - x^2y^2 + 2x^3y + 2x + 2$ as a sos using the connection to psd matrices and SDP.

11. Prove that a univariate non-negative polynomial is always a sum of two squares. (Hint: Make an argument about the possible real and complex roots of this polynomial and use the identity $(a^2 + b^2)(c^2 + d^2) = (ac + bd)^2 + (ad - bc)^2$ for all $a, b, c, d \in \mathbb{R}$.)

12. (Ex 3.35) Can you express $x^4 + 4x^3 + 6x^2 + 4x + 5$ as a sum of two squares?

13. (Ex 3.54) Let $p(x) = \sum_{k=0}^{2d} c_k x^k$. Give an explicit SDP formulation to compute the value of the global min of $p(x)$.

   (a) Show that the min of $p(x) = x^4 - 20x^2 + x$ is less than or equal to -100.

   (b) Show that the min of $p(x) = x^4 - 20x^2 + x$ is greater than or equal to -104.

   (c) Minimize the polynomial $p(x) = x^4 - 20x^2 + x$.

14. (a) (Ex 3.57) Find the value of p_{sos} for the polynomial $p(x,y,z) = x^4 + y^4 + z^4 - 4xyz + 2x + 3y + 4z$ over $\mathbb{R}^3$. Is $p_* = p^{\text{sos}}$ in this example? Do you expect $p_* = p^{\text{sos}}$?

    (b) Find the value of p_{sos} and p_* for the polynomial $p(x,y) = x^4 + y^4 - 4xyz$ over $\mathbb{R}^2$. Do you expect $p_* = p^{\text{sos}}$?

15. (Ex 3.69) Consider the quartic form in four variables:

$$p(w,x,y,z) = w^4 + x^2y^2 + x^2z^2 + y^2z^2 - 4wxyz.$$

    (a) Show that p is not a sos. (Hint: Use Reznick's result mentioned in Exercise 7).

    (b) Find a multiplier that makes the product a sos.

16. The Newton polytope of a polynomial $p(x_1,\dots,x_n)$ is the convex hull of all the non-negative integer vectors in $\mathbb{N}^n$ that appear as exponents of the monomials present in p. We will denote it as $\mathcal{N}(p)$. For example, $\mathcal{N}(x^2+xy+y^2)$ is the line segment in $\mathbb{R}^2$ that is the convex hull of $(2,0),(1,1),(0,2)$. Reznick proved the following theorem:

$$\text{If } p = \sum q_j^2 \text{ then } \mathcal{N}(q_i) \subseteq \tfrac{1}{2}\mathcal{N}(p) \text{ for each } i.$$

    (Ex 3.97)

    (a) Compute the Newton polytope of the Motzkin polynomial.

    (b) Which monomials would appear in a hypothetical sos decomposition of the Motzkin polynomial if you know the above theorem?

    (c) Show by considering the coefficient of x^2y^2, and the above calculation, that the Motzkin polynomial is not a sos.