Polynomial Optimization – Exercises

Rekha R. Thomas

September 10, 2016

Many of the exercises below are taken from the book *Semidefinite Optimization and Convex Algebraic Geometry* edited by Greg Blekherman, Pablo Parrilo and myself. The exercise number from the book is listed. A free pdf of this book is available at *http://www.math.washington.edu/~thomas/booksjournals.html* . Chapters 2-4 of this book are central to this course. Some of the examples also come from the monograph by Monique Laurent that you have a link to. Both of these sources contain a wealth of information on this subject.

If you see errors in these exercises, please send me an email (rrthomas@uw.edu).

1 Monday Lectures

1. Let $G = ([n], E)$ be an undirected graph where $[n] = \{1, \dots, n\}$ for a positive integer n. A set $S \subseteq [n]$ is said to be *stable* or *independent* if for every pair $i, j \in S$, 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.

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.

   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) 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) *Caratheodory's theorem* from convex geometry says that every element in a cone of dimension k can be written as a non-negative combination of at most k extreme rays of the cone. Can you prove this theorem?

   (d) Both the above exercises allow 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?

5. 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 1.1. A spectrahedron is a set of the form

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

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

   In the lecture we defined a *spectrahedron* as 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 $\operatorname{span}(A_1, \ldots, A_m)$ by A_0. If the matrices $A_1, \ldots, A_m$ are linearly independent in $\mathcal{S}^n$ then there is a bijection between the two versions in $\mathbb{R}^m$ and $\mathcal{S}^n$ respectively.

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

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

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

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

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

   (c) Consider the following concrete spectrahedron:

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

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 1.1.

6. 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?

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

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

   (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 $p \geq 0$ (resp. sos) $\;\Leftrightarrow\; \tilde{p} \geq 0$ resp. sos.

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

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

2 Tuesday Lectures

1. 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 $a, b, c, d \in \mathbb{R}$.)

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

3. (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 $p(x)$. Apply your formulation to the polynomial $p(x) = x^4 - 20x^2 + x$.

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

5. (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.

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

6. Polya's theorem states the following. Given a form $f(x_1, \ldots, x_n)$ that is strictly positive on the positive orthant, i.e., whenever $x_i \geq 0$ and $\sum x_i > 0$, then f can be expressed as $f = \frac{g}{h}$ where g and h are forms with positive coefficients. In particular, we can choose $h = (x_1 + x_2 + \cdots + x_n)^r$ for some r.

   (a) Argue the Reznick's theorem can be seen as a generalization of Polya's theorem.

   (b) (Ex 3.70) Consider the quadratic form $f(x, y) = (x-y)^2 + \epsilon xy$ which is obviously positive on the non-negative orthant for all $\epsilon > 0$. Estimate how large the exponent r must be, as a function of ϵ, for $(x+y)^r f(x, y)$ to have only positive coefficients.

7. The Newton polytope of a polynomial $p(x_1, \ldots, x_n)$ 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 RR^2 that is the convex hull of $(2, 0), (1, 1), (0, 2)$. Reznick proved the following theorem:

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

   (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.

8. We now examine an engineering application of sos polynomials that arises from dynamical systems and control theory. For more details see Sections 2.2.1 and 3.6.2 on the book *Semidefinite Optimization and Convex Algebraic Geometry.*

   Suppose we are given a system of differential equations

   $$\dot{x}_i(t) = f_i(x(t)), \ \ i = 1, \dots, n$$

   where f_i are polynomials in $x_1, \dots, x_n$. This system is globally asymptotically stable if there is an *energy function* $V(x) = V(x_1(t), \dots, x_n(t))$, called a *Lyapunov function*, such that

   $$V(x) > 0, \ \ \left(\frac{\partial V}{\partial x}\right)^{\top} f(x) < 0.$$

   Suppose we assume that $V(x)$ is a polynomial, then we may relax the nonnegativity requirement above to a sos condition as follows:

   $$V(x) \text{ is sos}, \quad -\left(\frac{\partial V}{\partial x}\right)^{\top} f(x) \text{ is sos}.$$

   Then we can solve for such a V by increasing the degree and using SDP.

   (Ex 3.173) Consider the polynomial dynamical system

   $$\begin{aligned} \dot{x} &= -x + (1+x)y \\ \dot{y} &= -(1+x)x \end{aligned}$$

   Find a polynomial Lyapunov function of degree four that proves global asymptotic stability of this system.

3 Thursday Lectures

1. Find $p_* = \inf\{10 - x^2 - y \,:\, x^2 + y^2 \le 1\}$. (It's easy to do some basic calculus to determine p_* in this example. You can use that to check the answer you get from the sos relaxation.)

2. Suppose we want to minimize a polynomial over an algebraic variety (given by equations) as opposed to a semialgebraic set:

$$p_* = \inf\{p(x) \,:\, g_1(x) = 0, \dots, g_m(x) = 0\}.$$

   (a) Write down the form of the p_t^{sos} problem in this case by modifying from a semialgebraic set to an algebraic set. What simplifications can you make?

   (b) Is there a way we can write a version of p_t^{sos} that is indifferent to the particular choice of equations defining the variety?

   (c) (Ex 3.99) Use your method to minimize the polynomial $10 - x^2 - y$ over the unit circle $x^2 + y^2 = 1$.

3. (Ex 3.62) Calculate p_1^{sos} for

$$\inf\{x^4 - 3x^2 + 1 \,:\, x^3 - 4x = 1\}.$$

4. Consider a system of polynomials $\{f_i(x) = 0 \; i = 1, \dots, m\}$ where $f_i \in \mathbb{R}[x]$.

   Hilbert's Nullstellensatz says that this system is infeasible over $\mathbb{C}^n$ if and only if -1 belongs to the ideal $\langle f_1, \dots, f_m\rangle$, i.e., there exists $F(x) \in \langle f_1, \dots, f_m\rangle$ such that $-1 = F(x)$.

   The *real Nullstellensatz* says that the system is infeasible over $\mathbb{R}^n$ if and only if -1 is congruent to a sos modulo the ideal $\langle f_1, \dots, f_m\rangle$, i.e., there exists $F(x) \in \langle f_1, \dots, f_m\rangle$ and a sos s such that $-1 = s + F(x)$.

   (Ex 3.135) Consider the set of equations:

$$\sum_{i=1}^{n} x_i = 1, \;\; x_i^2 = 0 \;\forall\; i = 1, \dots, n.$$

   (a) Check that this system is infeasible both over $\mathbb{R}$ and $\mathbb{C}$.

   (b) Give a real Nulstellensatz proof of infeasibility of this system over $\mathbb{R}$.

   (c) Show that every (Hilbert) Nullstellensatz proof of infeasibility must have degree greater than of equal to n.

5. (Ex 3.132) Consider the polynomial system

$$x + y^3 = 2, x^2 + y^2 = 1.$$

(a) How many complex solutions does the above system have?

(b) Does the system have any real solutions? If not find a Positivstellensatz based infeasibility certificate for this fact.

6. (Ex 3.134) Prove, using the alternative form of the Positivstellensatz, that every non-negative polynomial is a sum of squares of rational functions. (Hint: A polynomial $p(x) \geq 0$ for all $x \in \mathbb{R}^n$ if and only if the system $\{(x,y) \in \mathbb{R}^n \times \mathbb{R} : p(x) \leq 0,\ y \cdot p(x) = 1\}$ is empty.)

7. (a) Compare the Putinar and Schmudgen methods to prove that $x \leq -1$ and $x \leq 0$ on the unit disc with center at $(1,0)$ in the plane.

   (b) Repeat the same exercise over the compact region $\{(x,y) \in \mathbb{R}^2 : x^4 - x^3 - y^2 \leq 0\}$. It is helpful to draw the feasible region and observe its shape around the orgin where $x = 0$ is tangential.

8. Suppose $\{g_1(x) = \cdots = g_m(x) = 0\}$ is a polynomial system with only only finitely many complex (and hence also real) solutions. Also, assume that the ideal $\langle g_1, \ldots, g_m \rangle$ is radical (i.e., for every f^k in the ideal $\langle g_1, \ldots, g_m \rangle$, the polynomial f is also in the ideal).

   (a) Prove that whenever a polynomial p is nonnegative over the (real part of the) variety, then it is in fact a sos modulo the ideal. i.e., $p = \text{sos} + g$ where $g \in \langle g_1, \ldots, g_m \rangle$.

   (b) What can you say about the convergence of the Putinar hierarchy for the real variety defined by the g_i polynomials?

   (c) If $g_i = x_i^2 - x_i$ for all $i = 1, \ldots, n$, then its variety (both over $\mathbb{C}$ and $\mathbb{R}$) is the discrete hypercube $\{0,1\}^n$. Suppose we wanted to minimize a polynomial p over this discrete hypercube. Is there an upper bound that depends only on n for the number of steps needed by the Putinar hierachy to converge to p_*?

4 Friday Lectures

In the next few exercises we see how strong duality can fail in SDPs.

1. Write down the primal and dual SDP pair using the diagram for cone LPs.
2. (Ex 2.14) Consider the (primal) SDP:

$$p_* = \inf \ \{X_{11} : X_{22} = 0, \ X_{11} + 2X_{23} = 1, \ X \succeq 0\}.$$

   (a) Compute p_* and an optimal solution for this SDP.
   (b) Write down the dual of the above SDP.
   (c) Compute d^* and an optimal solution of the dual SDP if possible.
   (d) What do you conclude about this primal-dual pair?

3. (Ex 2.27) Consider the SDP:

$$d^* = \sup \left\{ y : \begin{bmatrix} 0 & y \\ y & 0 \end{bmatrix} \preceq \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\}.$$

   (a) What is d^*? Does this problem have an optimal solution?
   (b) Write the primal SDP.
   (c) Compute p_*. Does the primal problem have an optimum?

   In the following exercises we will explore conditions for strong duality.

4. (a) (Ex 2.30) We first note that a linear image of a closed cone may not be closed. Consider the proper cone $K = \{(x, y, z) \in \mathbb{R}^3 : y^2 \le xz, z \ge 0\}$. Show that K is a proper cone but its projection onto the x, y plane is not closed.
   (b) (Theorem 2.28) Let $K \subset V$ be a proper cone and $\mathcal{A} : V \to W$ be a linear map.Then show that the following two conditions are equivalent:
      i. $K \cap \ker(\mathcal{A}) = \{0\}$,
      ii. There exists $y \in W^*$ such that $\mathcal{A}^* y \in \text{int}(K^*)$.

      Prove that if these conditions hold, then $\mathcal{A}(K)$ is closed.
   (c) Closed cones allow us to find nice feasibility certificates. Suppose we want to decide the feasibility of the system

$$\mathcal{A}(x) = b, \ x \in K$$

      for a proper cone K and linear map $\mathcal{A}$. If $\mathcal{A}(K)$ is closed then show that the above system is infeasible if and only if there exists y such that

$$\langle y, b \rangle = -1, \ \mathcal{A}^* y \in K^*.$$

(Hint: If we have a compact set U and a closed set W in $\mathbb{R}^n$, then they are disjoint if and only if there is a hyperplane *separating* them, in the sense that U lies in one open half space of the hyperplane and W lies in the opposite closed halfspace.)

Note that if $K = \mathbb{R}^n_+$, then K is a polyhedron and $\mathcal{A}(K)$ is always closed and we will get infeasibility certificates of the above type for systems of the form $Mx = b, x \geq 0$ where M is a matrix.

(d) Consider a primal cone program

$$p_* = \inf\{\langle c, x\rangle : \mathcal{A}(x) = b,\ x \in K\},$$

where $\mathcal{A} : V_1 \to V_2$ is a linear map. Define the extended map:

$$\hat{\mathcal{A}} : V_1 \to V_2 \oplus \mathbb{R}, \quad x \mapsto (\mathcal{A}(x), \langle c, x\rangle).$$

If $\hat{\mathcal{A}}(K)$ is closed and the above primal problem is feasible, then strong duality holds, i.e., $p_* = d^*$ and the primal has an optimum. Prove this theorem or see (Barvinok Chapter IV, 7.2) for a proof.

(e) Using the above results argue that the following holds for SDPs: If D is strictly feasible and P is feasible then $p_* = d^*$ and P has an optimum. In particular, if both the primal and dual are strictly feasible then both problems have optima and their optimal values coincide. (Recall that strict feasibility means that there is a *positive definite* matrix in the feasible region – one with all eigenvalues strictly positive.)

5. Write down the p^{mom} problem for minimizing the Motzkin polynomial as we did in the lecture. Recall that

$$p = x_1^4 x_2^2 + x_1^2 x_2^4 - 3x_1^2 x_2^2 + 1.$$

Can you argue that $p^{\mathrm{mom}} = -\infty$? Recall that $p^{\mathrm{sos}} = -\infty$.