Discussion 14

1. Sum of Squares

Given a polynomial $p(t)$ in a single variable t, we are interested in knowing if we can write $p(t)$ as a sum of squares of polynomials, i.e. whether we can write

$$p(t) = \sum_{j=1}^{k} (q_j(t))^2,$$

for some $k \geq 1$ and some polynomials $q_1(t), \cdots, q_k(t)$. This is an interesting question in many contexts, because if we could do this then we would know that $p(t)$ is nonnegative for all values of t. [1]

First observe that if $p(t)$ can be written as a sum of squares of polynomials then it must have even degree. We will therefore assume that $p(t)$ has degree $2d$ for some integer $d \geq 1$ (the case $d = 0$ corresponds to $p(t)$ being a constant).

Let $\vec{z} := \begin{bmatrix} 1 & t & \cdots & t^d \end{bmatrix}^\top$. Note that $\vec{z}$ is a $(d+1)$-dimensional vector whose entries are polynomials in t.

(a) Show that the polynomial $p(t)$ of degree $2d$ can be written as a sum of squares of polynomials if and only if there is a symmetric positive semidefinite matrix Q such that

$$p(t) = \vec{z}^\top Q \vec{z}.$$

(The equality here is an equality between polynomials in t.)

Hint: Every symmetric positive semidefinite matrix Q can be written as a sum of dyads, i.e. $Q = \sum_{i=1}^{n} \vec{u}_i \vec{u}_i^\top$ if $Q \in \mathbb{S}_+^n$.

Solution: Suppose we can write

$$p(t) = \sum_{j=1}^{k} (q_j(t))^2,$$

for some $k \geq 1$. Note that each $q_j(t)$ must have degree at most d. Let $\vec{u}_j \in \mathbb{R}^{d+1}$ denote the vector of its coefficients, in the sense that

$$q_j(t) = \vec{u}_j^\top \vec{z},\ j = 1, \cdots, k,$$

or equivalently,

$$q_j(t) = u_{j1} + u_{j2} t + \cdots + u_{j(d+1)} t^d,\ j = 1, \cdots, k,$$

where $\vec{u}_j = \begin{bmatrix} u_{j1} & \cdots & u_{j(d+1)} \end{bmatrix}^\top \in \mathbb{R}^{d+1}$. Then we have

$$p(t) = \sum_{j=1}^{k} (q_j(t))^2 = \sum_{j=1}^{k} (\vec{u}_j^\top \vec{z})^2 = \vec{z}^\top \left(\sum_{j=1}^{k} \vec{u}_j \vec{u}_j^\top \right) \vec{z}.$$

Defining $Q := \sum_{j=1}^{k} \vec{u}_j \vec{u}_j^\top$, note that Q is symmetric positive semidefinite and we have $p(t) = \vec{z}^\top Q \vec{z}$.

Conversely, if we could write $p(t) = \vec{z}^\top Q \vec{z}$ for some symmetric positive semidefinite matrix Q, then, since we can write Q as a sum of k dyads, where $k = \text{rank}(Q)$, i.e. $Q = \sum_{j=1}^{k} \vec{u}_j \vec{u}_j^\top$ for some $\vec{u}_j \in \mathbb{R}^{d+1}$, $1 \leq j \leq k$, we would have

$$p(t) = \vec{z}^\top \left(\sum_{j=1}^{k} \vec{u}_j \vec{u}_j^\top \right) \vec{z} = \sum_{j=1}^{k} (\vec{u}_j^\top \vec{z})^2 = \sum_{j=1}^{k} (q_j(t))^2,$$

[1] In fact, it is known that if a polynomial $p(t)$ in a single variable is nonnegative for all values of t then it can be written as a sum of squares of two polynomials, i.e. $p(t) = r(t)^2 + s(t)^2$ for some polynomials $r(t)$ and $s(t)$, but we do not need this fact.

where we define the polynomial $q_j(t) := \vec{u}_j^\top \vec{z}$ for $1 \leq j \leq k$. This shows that $p(t)$ can be written as a sum of squares of polynomials. We have thus shown that $p(t)$ can be written as a sum of squares of polynomials if and only if we can write $p(t) = \vec{z}^\top Q \vec{z}$ for some symmetric positive semidefinite matrix Q.

(b) Show that we can pose the question of whether a given polynomial $p(t)$ of degree $2d$ can be written as a sum of squares of polynomials as a feasibility question for an SDP in standard form.

Remark: Recall that an SDP in standard form looks like:

$$\begin{aligned} \min_{X \in \mathbb{R}^{n \times n}} \quad & \text{trace}(CX) \\ \text{s.t.} \quad & \text{trace}(A_i X) = b_i, \text{ for each } i \in \{1, \cdots, m\}, \\ & X \succeq 0. \end{aligned}$$

Here the minimization is over matrices $X \in \mathbb{S}^n$. The matrices $C, A_1, \cdots, A_m \in \mathbb{S}^n$ as well as the scalars $b_1, \cdots, b_m \in \mathbb{R}$ are given. The constraint $X \succeq 0$ is the constraint that X should be symmetric positive semidefinite.

Also recall that to pose a minimization problem as a feasibility problem, we can just take the objective to be the constant 0 (so the question then just becomes whether the value of the problem is 0, in which case the problem is feasible, or ∞, in which case the problem is infeasible). For an SDP in standard form to be a feasibility problem, therefore, we could just take C to be the zero matrix.

Solution: Suppose

$$p(t) = b_1 + b_2 t + \cdots + b_{2d+1} t^{2d},$$

where $b_1, b_2, \cdots, b_{2d+1} \in \mathbb{R}$, and suppose

$$Q = \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1(d+1)} \\ q_{21} & q_{22} & & q_{2(d+1)} \\ \vdots & & & \vdots \\ q_{(d+1)1} & q_{(d+1)2} & \cdots & q_{(d+1)(d+1)} \end{bmatrix}.$$

The equation

$$p(t) = \vec{z}^\top Q \vec{z}$$

is equivalent to the equations

$$\text{trace}(A_i Q) = b_i, \text{ for each } i \in \{1, \cdots, 2d+1\},$$

where $A_i \in \mathbb{R}^{(d+1)\times(d+1)}$ is the matrix with zeros everywhere except for the coordinates (u, v) that satisfy $u + v = i + 1$.

The question of whether $p(t)$ can be written as a sum of squares of polynomials can therefore be posed as the following SDP feasibility problem in standard form:

$$\begin{aligned} \min_{X} \quad & 0 \\ \text{s.t.} \quad & \text{trace}(A_i X) = b_i, \text{ for each } i \in \{1, \cdots, 2d+1\}, \\ & X \succeq 0. \end{aligned}$$

Indeed, this problem is feasible if and only if there is a symmetric positive semidefinite Q such that $p(t) = \vec{z}^\top Q \vec{z}$.

2. SDP Duality

Consider the following SDP in inequality form:

$$\begin{aligned} \min_{(x,y)\in\mathbb{R}^2} \quad & x \\ \text{s.t.} \quad & \begin{bmatrix} x & 1 \\ 1 & y \end{bmatrix} \succeq 0. \end{aligned} \tag{1}$$

(a) Draw the feasible set. Is it convex?

Solution: The symmetric matrix $\begin{bmatrix} x & 1 \\ 1 & y \end{bmatrix}$ is symmetric positive semidefinite if and only if its diagonal entries and its determinant are nonnegative, i.e. $x \geq 0$, $y \geq 0$ and $xy \geq 1$. The region of (x, y) defined by these conditions is the epigraph of the convex function $y = \frac{1}{x}$ with domain $\mathbb{R}_{++} := \{(x, y) \in \mathbb{R}^2 | x, y > 0\}$. Being the epigraph of a convex function, this set is a convex subset of $\mathbb{R}^2$. See Figure 1.

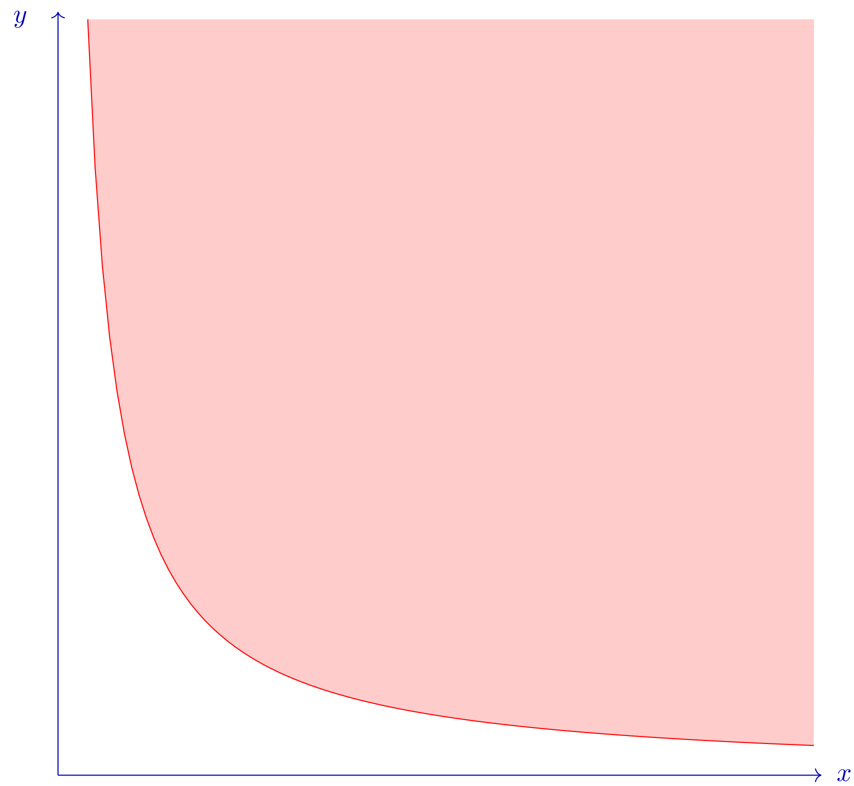

Figure 1: The domain of the primal SDP is the closed region above the curve $xy = 1$ in the nonnegative quadrant. This is a convex set.

(b) Write the conic dual SDP in standard form.

Solution: If the primal SDP in inequality form is given by

$$\begin{aligned} \min_{\vec{x}\in\mathbb{R}^m} \quad & \vec{c}^\top \vec{x} \\ \text{s.t.} \quad & F_0 + \sum_{i=1}^{m} x_i F_i \succeq 0, \end{aligned}$$

where $F_0, F_1, \cdots, F_m \in \mathbb{S}^n$ and $c \in \mathbb{R}^n$, its dual is the SDP in standard form given by

$$\max_{X\in\mathbb{S}^n} \ \text{trace}(-F_0 X)$$

$$\text{s.t. } \operatorname{trace}(F_i X) = c_i, \text{ for each } i \in \{1, \cdots, n\},$$
$$X \succeq 0.$$

Here we have $n = 2$, with the vectors in $\mathbb{R}^2$ written as $\begin{bmatrix} x & y \end{bmatrix}^\top$, and

$$c = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \; F_0 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \; F_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \; F_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

The dual of the primal SDP in (1) is given by

$$\begin{aligned} \max_{x_{11}, x_{12}, x_{22}} \quad & -2x_{12} \\ \text{s.t. } \quad & x_{11} = 1, \\ & x_{22} = 0, \\ & \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix} \succeq 0. \end{aligned} \tag{2}$$

To make it clear that it is in standard form, the dual of the primal SDP can be written explicitly as below if desired, although writing it in a more straightforward way as above would also be fine:

$$\begin{aligned} \max_{X \in \mathbb{S}^2} \quad & \operatorname{trace}\left(\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} X \right) \\ \text{s.t.} \quad & \operatorname{trace}\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} X \right) = 1, \\ & \operatorname{trace}\left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} X \right) = 0, \\ & X \succeq 0. \end{aligned}$$

(c) Is the primal SDP feasible? Is it strictly feasible?

Remark: The SDP in inequality form

$$\begin{aligned} \min_{\vec{x} \in \mathbb{R}^m} \quad & \vec{c}^\top \vec{x} \\ \text{s.t. } \quad & F_0 + \sum_{i=1}^m x_i F_i \succeq 0, \end{aligned}$$

where $F_0, F_1, \cdots, F_m \in \mathbb{S}^n$, $\vec{c} \in \mathbb{R}^n$, is said to be strictly feasible if there is some $\vec{x} \in \mathbb{R}^n$ such that $F(\vec{x}) \in \mathbb{S}^n_{++}$, i.e. $F(\vec{x})$ is symmetric positive definite. Here $F(\vec{x})$ denotes $F_0 + \sum_{i=1}^m x_i F_i$.

Solution: Since we can choose x and y such that $\begin{bmatrix} x & 1 \\ 1 & y \end{bmatrix}$ is symmetric positive definite, the primal SDP in (1) is strictly feasible (and hence also feasible).

(d) Is the dual SDP feasible? Is it strictly feasible?

Remark: The SDP in standard form

$$\min_{X \in \mathbb{S}^n} \; \operatorname{trace}(CX)$$

$$\text{s.t. } \operatorname{trace}(A_i X) = b_i, \text{ for each } i \in \{1, \cdots, m\},$$
$$X \succeq 0,$$

where $C, A_1, \cdots, A_m \in \mathbb{S}^m$, $b_1, \cdots, b_m \in \mathbb{R}$, is said to be strictly feasible if there is some $X \in \mathbb{S}^n_{++}$ (symmetric positive definite) satisfying the equality constraints $\operatorname{trace}(A_i X) = b_i$ for each $i \in \{1, \cdots, m\}$.

Solution: From the dual SDP given in (2), we see that since we must have $x_{22} = 0$, the condition $\begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix} \succeq 0$ forces $x_{12} = 0$. Since we also require that $x_{11} = 1$, there is only one matrix X in the feasible set, namely $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. This matrix is not symmetric positive definite. We conclude that the dual SDP is feasible but not strictly feasible.

(e) Find the optimal primal value p^* and the optimal dual value d^*. Does strong duality hold?

Solution: From Figure 1 it is clear that the optimal primal value is $p^* = 0$ (which is not attained, i.e. there is no optimal point). To see this analytically, first note that we must have $x \geq 0$ at every feasible point. Further, for any $\epsilon > 0$, however small, the matrix $\begin{bmatrix} \epsilon & 1 \\ 1 & \epsilon^{-1} \end{bmatrix}$ is symmetric positive semidefinite, so the objective function of the primal problem can be made to equal ϵ.

As for the dual problem, it has only one feasible point and the value of the dual objective at that feasible point is 0, so we have $d^* = 0$.

Since $p^* = d^*$, strong duality holds.

