EECS 127/227AT **Fall 2020**
Discussion 10

1 Generalized geometric programming

A *geometric program* (GP) takes the form

$$\begin{aligned} \min_{x_1,\ldots,x_n} \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \le 1, \; i = 1, \ldots, m, \\ & h_j(x) = 1, \; j = 1, \ldots, p, \end{aligned} \tag{1}$$

where $f_i(x)$ for $0 \le i \le m$ are *posynomials* and $h_j(x)$ for $1 \le j \le p$ are *monomials*. Here $x = (x_1, \ldots, x_n)$, and we recall that a function of the form $\delta x_1^{d_1} \ldots x_n^{d_n}$ in the positive variables $x_1, \ldots, x_n$, with $d_i \in \mathbb{R}$ for $1 \le i \le n$ and $\delta > 0$ is called a monomial in the theory of geometric programming. The domain of the monomial is $\mathbb{R}^n_{++}$. A posynomial, by definition, is a finite sum of monomials.

Every GP can be rewritten as a convex optimization problem by replacing the objective function $f_0(x)$ by $\log f_0(e^y)$, replacing the inequalities $f_i(x) \le 1$ by the inequalities $\log f_i(e^y) \le 0$, and replacing the equality conditions $h_j(x) = 1$ by the equality conditions $\log h_j(e^y) = 0$. Here $y = (y_1, \ldots, y_n)$, with x_i being replaced by e^{y_i} for $1 \le i \le n$.

A *generalized geometric program* (GGP) takes the form

$$\begin{aligned} \min_{x_1,\ldots,x_n} \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \le 1, \; i = 1, \ldots, m, \\ & h_j(x) = 1, \; j = 1, \ldots, p, \end{aligned} \tag{2}$$

where $f_i(x)$ for $0 \le i \le m$ are *generalized posynomials* and $h_j(x)$ for $1 \le j \le p$ are monomials. A function of the positive variables $x_1, \ldots, x_n$ is called a generalized posynomial if it can be formed from posynomials by the operations of addition, multiplication, taking positive (fractional) powers and taking maximums.

It turns out that every GGP can be converted into a GP by introducing new variables. GGP solvers will typically do this as part of the process of expressing a GGP as a convex optimization problem before solving it. In this question we will illustrate via an example how a GGP can be rewritten as a GP.

Consider the optimization problem

$$\begin{aligned} \min_{x,y,z} \quad & \max(x, y) \\ \text{s.t.} \quad & x^2 + y \le \sqrt{xyz}, \\ & \max(y, z) \le \frac{1}{\sqrt{x+z}}, \\ & xyz = 1. \end{aligned} \tag{3}$$

(a) Is the optimization problem in (3) a geometric program?

(b) Is the optimization problem in (3) a generalized geometric program?

(c) Rewrite the GGP in (3) as a GP.

2 SDPs and congruence transformations

Consider the SDP

$$\begin{aligned} &\min_x \; c^\top x \\ &\text{s.t.}: \; x_1 F_1 + \ldots + x_n F_n + G \succeq 0. \end{aligned} \tag{4}$$

Here $G, F_1, \ldots, F_n \in \mathbb{S}^k$ and $c \in \mathbb{R}^n$.
Let $R \in \mathbb{R}^{k \times k}$ be a nonsingular matrix. Define

$$\begin{aligned} \tilde{G} &= R^\top G R, \\ \tilde{F}_i &= R^\top F_i R, \; i = 1, \ldots, n, \end{aligned}$$

$\tilde{G}$ is a congruence transformation of G and the $\tilde{F}_i$ are congruence transformations of the respective F_i.

(a) Show that the SDP

$$\begin{aligned} &\min_x \; c^\top x \\ &\text{s.t.}: \; x_1 \tilde{F}_1 + \ldots + x_n \tilde{F}_n + \tilde{G} \succeq 0, \end{aligned} \tag{5}$$

is equivalent to the SDP in (4).

Hint: Both problems involve the minimization of the same objective function over the same variable, so we only need to show their feasible sets are the same.

(b) Suppose R is such that $\tilde{G}$ and each of the $\tilde{F}_i$ for $1 \leq i \leq n$ is a diagonal matrix. Show that the SDP in (4) is equivalent to an LP.

(c) Suppose R is such that we have

$$\begin{aligned} \tilde{G} &= \begin{bmatrix} \beta I_{k-1} & b \\ b^\top & \beta \end{bmatrix}, \\ \tilde{F}_i &= \begin{bmatrix} \alpha_i I_{k-1} & a_i \\ a_i^\top & \alpha_i \end{bmatrix}, \; i = 1, \ldots, n, \end{aligned}$$

for some $\beta \in \mathbb{R}$, $b \in \mathbb{R}^{k-1}$ and $\alpha_i \in \mathbb{R}$, $a_i \in \mathbb{R}^{k-1}$ for $1 \leq i \leq n$. Show that the SDP in (4) is equivalent to an SOCP.

3 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 $q_1(t), \ldots, q_k(t)$. 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] [2]

Let us first make the simple observation 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 $d = 0$ is trivial).

Let $z := \begin{bmatrix} 1 & t & \ldots & t^d \end{bmatrix}^\top$. Note that z is $d+1$-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 iff there is positive semidefinite matrix Q such that

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

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

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

(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 is standard form looks like:

$$\begin{aligned} \min_X \;\; & \text{trace}(CX) \\ \text{s.t.} \;\; & \text{trace}(A_i X) = b_i, \; i = 1, \ldots, m, \\ & X \succeq 0. \end{aligned}$$

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

[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.

[2] More generally, we can use this to get lower bounds for $p(t)$ that apply for all t. To check if $p(t) \geq c$ for all values of t, we just check if the polynomial $p(t) - c$ can be written as a sum of squares of polynomials.