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?

Solution:
No. For instance, the objective function $\max(x,y)$ is not a posynomial.

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

Solution:
Yes.

The objective function $\max(x, y)$ is the maximum of two posynomials (in fact it is the maximum of two monomials).

The inequality constraint $x^2+y \le \sqrt{xyz}$ can be written as $x^{1.5}y^{-0.5}z^{-0.5}+x^{-0.5}y^{0.5}z^{-0.5} \le 1$. Here $x^{1.5}y^{-0.5}z^{-0.5}+x^{-0.5}y^{0.5}z^{-0.5}$ is a posynomial (it would have been enough for it to be a generalized posynomial).

The inequality constraint $\max(y, z) \le \frac{1}{\sqrt{x+z}}$ can be written as $\max(y\sqrt{x+z}, z\sqrt{x+z}) \le 1$. Here $\max(y\sqrt{x+z}, z\sqrt{x+z})$ is a generalized polynomial, because it is the maximum of two generalized posynomials. To see this, note that $y\sqrt{x+z}$ is a generalized posynomial because it is the product of the monomial y and the generalized posynomial $\sqrt{x+z}$, where we see that $\sqrt{x+z}$ is a generalized posynomial because it is a positive fractional power of the posynomial $x+z$. The logic to show that $z\sqrt{x+z}$ is a generalized posynomial is similar.

In the equality constraint $xyz = 1$, we note that xyz is a monomial, as is required.

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

Solution:
To deal with the fact that the objective function is not a posynomial, we can introduce a new variable t to represent $\max(x, y)$. This will introduce two new inequalities $x \le t$ and $y \le t$, which can be written as $xt^{-1} \le 1$ and $yt^{-1} \le 1$ respectively, with xt^{-1} and yt^{-1} being posynomials (in fact they are monomials).

We could try a similar idea to deal with the second inequality, by introducing a variable s to represent $\sqrt{x+z}$ but we would run into trouble if we did this because the equality $s = \sqrt{x+z}$ cannot be expressed as a monomial being equal to 1, so is not allowed. We can get around this problem by instead thinking of s as representing a variable that is at least as big as $\sqrt{x+z}$. This leads to the following GP, which is equivalent to the GGP in (3)

$$
\begin{aligned}
\min_{x,y,z,s,t} \quad & t \\
\text{s.t.} \quad & x^2+y \le \sqrt{xyz}, \\
& x \le t, \\
& y \le t, \\
& x+z \le s^2, \\
& ys \le 1, \\
& zs \le 1, \\
& xyz = 1.
\end{aligned}
$$

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.

Solution:
$x \in \mathbb{R}^n$ lies in the feasible set of the SDP in (4) iff we have

$$u^\top \left(\sum_{i=1}^n x_i F_i + G\right) u \geq 0, \ \text{for all } u \in \mathbb{R}^k.$$

Since R is nonsingular this is equivalent to the condition

$$(Ru)^\top \left(\sum_{i=1}^n x_i F_i + G\right) Ru \geq 0, \ \text{for all } u \in \mathbb{R}^k,$$

which is the same as the condition

$$u^\top \left(\sum_{i=1}^n x_i \tilde{F}_i + G\right) u \geq 0, \ \text{for all } u \in \mathbb{R}^k,$$

but this, in turn, is precisely the condition for x to lie in the feasible set of the SDP in (5).

Since the feasible sets of the two SDPs are the same, as we have just established, and their objective functions are also the same, they are equivalent.

(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.

Solution:
Suppose $\tilde{G} = \mathrm{diag}(\tilde{g}_1, \ldots, \tilde{g}_k)$ and $\tilde{F}_i = \mathrm{diag}(\tilde{f}_{i1}, \ldots, \tilde{f}_{ik})$ for $1 \le i \le n$. Then $x \in \mathbb{R}^n$ belongs to the feasible set of the SDP in (5) iff we have

$$\sum_{i=1}^{n} x_i \tilde{f}_{ij} + \tilde{g}_j \ge 0, \text{ for all } 1 \le j \le k.$$

This is because a diagonal matrix is positive semidefinite iff all its diagonal entries are nonnegative.

We thus see that the SDP in (5) is actually an LP. Since we have established in the preceding part of the question that the SDP in (4) is equivalent to the one in (5), we see that the SDP in (4) is equivalent to an LP.

(c) Suppose R is such that we have

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

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 $1 \le i \le n$. Show that the SDP in (4) is equivalent to an SOCP.

Solution:
If R is such that $\tilde{G}$ and $\tilde{F}_i$ for $1 \le i \le n$ have the stated forms, we see that $x \in \mathbb{R}^n$ is in the feasible set of the SDP in (5) iff we have

$$\begin{bmatrix} (\alpha^\top x + \beta) I_{k-1} & Ax + b \\ (Ax + b)^\top & (\alpha^\top x + \beta) \end{bmatrix} \succeq 0,$$

where $\alpha \in \mathbb{R}^n$ denotes the column vector of the α_i and $A \in \mathbb{R}^{(k-1)\times n}$ denotes the matrix whose columns are the a_i.

For this to hold we must first of all have $\alpha^\top x + \beta \ge 0$, as we can see from the leading $(k-1)\times(k-1)$ block of this matrix.

If we have $\alpha^\top x + \beta = 0$ then we must have $Ax + b = 0$. To see this, suppose we have $\alpha^\top x + \beta = 0$ and $Ax + b \ne 0$. Then we must have $(Ax + b)_l = z \ne 0$ for some $1 \le l \le k-1$. If $z > 0$, let $v := -e_l + e_k$, while if $z < 0$ let $v := e_l - e_k$, where $e_1, \ldots, e_k \in \mathbb{R}^k$ denote the standard unit vectors. We would then have

$$v^\top \begin{bmatrix} (\alpha^\top x + \beta) I_{k-1} & Ax + b \\ (Ax + b)^\top & (\alpha^\top x + \beta) \end{bmatrix} v = -2|z| < 0,$$

which contradicts the desired positive semidefiniteness.

If $\alpha^\top x + \beta > 0$, then the Schur complement of the leading $(k-1)\times(k-1)$ block of this matrix must be positive semidefinite, i.e. we must have

$$(\alpha^\top x + \beta) - \frac{1}{\alpha^\top x + \beta}(Ax + b)^\top (Ax + b) \ge 0,$$

Putting these conditions together, we see that the condition for $x \in \mathbb{R}^n$ to be feasible for the SDP in (5) is

$$\|Ax + b\|_2 \le \alpha^\top x + \beta.$$

Hence the SDP in (5) is an SOCP. Since we have established in the first part of the question that the SDP in (4) is equivalent to the one in (5), we see 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 $Q \in \mathbb{S}_+^n$.

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 $u_j \in \mathbb{R}^{d+1}$ denote the vector of its coefficients, in the sense that

$$q_j(t) = u_j^\top z, \; j = 1, \ldots, k,$$

or equivalently,

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

where $u_j = \begin{bmatrix} u_{j1} & \ldots & u_{j(d+1)} \end{bmatrix}^\top$. Then we have

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

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

[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.

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

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

where we define the polynomial $q_j(t) := u_j^\top z$ for $1 \le j \le k$. 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 iff we can write $p(t) = z^\top Q z$ for some 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 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.

Solution:
Suppose

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

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

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

is equivalent to the equations

$$\text{trace}(A_i Q) = b_i, \ i = 1, \ldots, 2d+1,$$

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

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:

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

Indeed, this problem is feasible iff there is a positive semidefinite Q such that $p(t) = z^\top Q z$.