Math 484: Nonlinear Programming[1] Mikhail Lavrov

# Chapter 2, Lecture 7: Solving the dual geometric program

*February 18, 2019* *University of Illinois at Urbana-Champaign*

## 1 The geometric programming algorithm

Here is a summary of what we know about geometric programming so far. Given a geometric program, a problem of the form

$$\begin{aligned} \underset{\mathbf{t}\in\mathbb{R}^m}{\text{minimize}} \quad & g(\mathbf{t}) = \text{Term}_1(\mathbf{t}) + \text{Term}_2(\mathbf{t}) + \cdots + \text{Term}_n(\mathbf{t}) \\ \text{subject to} \quad & t_1, t_2, \dots, t_m > 0 \end{aligned}$$

where for $1 \le i \le n$, $\text{Term}_i(\mathbf{t})$ is a posynomial term

$$\text{Term}_i(\mathbf{t}) = C_i t_1^{\alpha_{i1}} t_2^{\alpha_{i2}} \cdots t_m^{\alpha_{im}},$$

we begin by constructing the dual geometric program

$$\begin{aligned} \underset{\boldsymbol{\delta}\in\mathbb{R}^n}{\text{maximize}} \quad & v(\boldsymbol{\delta}) = \left(\frac{C_1}{\delta_1}\right)^{\delta_1} \left(\frac{C_2}{\delta_2}\right)^{\delta_2} \cdots \left(\frac{C_n}{\delta_n}\right)^{\delta_n} \\ \text{subject to} \quad & \delta_1\alpha_{11} + \delta_2\alpha_{21} + \cdots + \delta_n\alpha_{n1} = 0, && \text{(power of } t_1) \\ & \delta_1\alpha_{12} + \delta_2\alpha_{22} + \cdots + \delta_n\alpha_{n2} = 0, && \text{(power of } t_2) \\ & \dots \\ & \delta_1\alpha_{1m} + \delta_2\alpha_{2m} + \cdots + \delta_n\alpha_{nm} = 0, && \text{(power of } t_m) \\ & \delta_1 + \delta_2 + \cdots + \delta_n = 1, && \text{(AM-GM requirement)} \\ & \delta_1, \delta_2, \dots, \delta_n > 0. && \text{(AM-GM requirement)} \end{aligned}$$

and solving it to get the optimal dual solution $\boldsymbol{\delta}^*$. Next, using this dual solution, we find $\mathbf{t}^*$ by solving the equations

$$\frac{\text{Term}_1(\mathbf{t}^*)}{\delta_1^*} = \frac{\text{Term}_2(\mathbf{t}^*)}{\delta_2^*} = \cdots = \frac{\text{Term}_n(\mathbf{t}^*)}{\delta_n^*} = v(\boldsymbol{\delta}^*).$$

The result is guaranteed to be an optimal solution to the original problem.

## 2 Duality in optimization

This dual-problem approach is an idea that occurs in many places in optimization. If you've taken a class on linear programming, you've seen dual linear programs; later in this course, we'll see that both linear programming duality and geometric programming duality are special cases of KKT duality.

[1] This document comes from the Math 484 course webpage: https://faculty.math.illinois.edu/~mlavrov/courses/484-spring-2019.html

But even more broadly than this, the philosophy of optimization duality is as follows: we suppose that we have an optimization problem and a scheme for coming up with bounds on it. (In the case of a minimization problem, they should be lower bounds.) Our scheme should be flexible: there should be some parameters we can vary that determine how good a bound we get. In this case, the dual problem is to choose these parameters to find the best bound possible.

To talk about this more easily, here is some terminology:

- The set of all points satisfying the constraints of an optimization problem is called the *feasible region*, and a point in that region is a *feasible solution* or *candidate solution*.
- The function we are trying to optimize is called the *objective function*, and a feasible solution with the best objective value possible is an *optimal solution*.
- When we have a dual program, the original problem is called the *primal program*, and we talk about points that are *primal feasible*, *dual optimal*, etc.
- When a problem has no feasible solutions, we call it *infeasible* or *inconsistent*.

Duality is symmetric. For example, instead of viewing a dual feasible solution $\boldsymbol{\delta}$ as a parameter giving us a lower bound $v(\boldsymbol{\delta})$ on the primal objective value, we can think of a primal feasible solution $\mathbf{t}$ as a parameter giving us an upper bound $g(\mathbf{t})$ on the dual objective value. Both of these are ways to state the inequality that for every primal feasible $\mathbf{t}$ and dual feasible $\boldsymbol{\delta}$,

$$g(\mathbf{t}) \geq v(\boldsymbol{\delta}).$$

So the primal program is a dual of the dual program.

In the case of geometric programming, not only do we get the inequality $g(\mathbf{t}) \geq v(\boldsymbol{\delta})$ (which is a required feature of optimization duality), but also something stronger: if $\mathbf{t}^*$ is the optimal primal solution and $\boldsymbol{\delta}^*$ is the optimal dual solution, then $g(\mathbf{t}^*) = v(\boldsymbol{\delta}^*)$. This is called *strong duality*.

In other cases, there might be a *duality gap* between the optimal values of the primal and the dual. The dual program is doing the best it can to give us bounds on the primal program, but the best bound it can give is still not tight.

## 3 Badly behaved geometric programs

To be able to solve geometric programs properly, we have to know what can go wrong, and be able to recognize when we are dealing with a badly behaved geometric program.

For now—while our geometric programs are still unconstrained—we are lucky that they are always consistent. Set all variables to 1, and you will get a feasible solution. Still, there can be trouble: the geometric program

$$\begin{aligned} \underset{x \in \mathbb{R}}{\text{minimize}} \quad & g(x) = x \\ \text{subject to} \quad & x > 0 \end{aligned}$$

has no optimal solution, because we want to set $x$ to 0 and we're not allowed to do that.

This is reflected in the dual program. The dual will have exactly one variable $\delta$ in this case. We get two conflicting requirements: $\delta = 0$ so that the power of $x$ is 0, and $\delta = 1$ so that we can apply the AM-GM inequality. So the dual program is inconsistent.

Here's a slightly fancier version of this bad behavior. Consider the geometric program

$$\begin{aligned} \underset{x,y\in\mathbb{R}}{\text{minimize}} \quad & g(x,y) = x + x^{-1} + xy^{-1} \\ \text{subject to} \quad & x > 0, y > 0. \end{aligned}$$

Here, we can make $g(x,y)$ arbitrarily close to 2 by setting $x = 1$ and setting $y$ to a very large number. But we can never get to 2, because $y^{-1} > 0$ for all $y > 0$. So, once again, there is no optimal solution.

The dual geometric program is

$$\begin{aligned} \underset{\boldsymbol{\delta}\in\mathbb{R}^3}{\text{maximize}} \quad & v(\boldsymbol{\delta}) = \left(\frac{1}{\delta_1}\right)^{\delta_1} \left(\frac{1}{\delta_2}\right)^{\delta_2} \left(\frac{1}{\delta_3}\right)^{\delta_3} \\ \text{subject to} \quad & \delta_1 - \delta_2 + \delta_3 = 0, && \text{(power of } x\text{)} \\ & -\delta_3 = 0, && \text{(power of } y\text{)} \\ & \delta_1 + \delta_2 + \delta_3 = 1, && \text{(AM-GM requirement)} \\ & \delta_1, \delta_2, \delta_3 > 0. && \text{(AM-GM requirement)} \end{aligned}$$

This is inconsistent for a subtler reason: solving the linear equations gives us the point $\boldsymbol{\delta} = (\frac{1}{2}, \frac{1}{2}, 0)$, but this violates the positivity constraint $\delta_3 > 0$.

We will not prove this yet, but these examples are instances of a general principle: the dual geometric program is inconsistent precisely when the primal program has no optimal solution. So when the primal program is ill-behaved, we'll catch it.

## 4 An example with infinitely many dual feasible solutions

Here is an example of what happens when there are infinitely many feasible dual solutions, so that we actually have to consider $v(\boldsymbol{\delta})$ when solving the dual program. (In general, we expect this to happen when $g(\mathbf{t})$ has few variables and many terms, since this corresponds to a dual program with many variables and few constraints.)

The problem is this:

$$\begin{aligned} \underset{t_1,t_2\in\mathbb{R}}{\text{minimize}} \quad & g(t_1,t_2) = t_1^2 + t_2^2 + 2t_1t_2 + \frac{1}{t_1t_2} \\ \text{subject to} \quad & t_1, t_2 > 0. \end{aligned}$$

The dual problem will have four variables $\delta_1, \delta_2, \delta_3, \delta_4$, corresponding to the four terms. We get three constraints, one from $t_1$, one from $t_2$, and one from the usual requirement that the $\delta$'s all have to sum to 1.

$$
\begin{array}{lll}
\underset{\delta_1,\delta_2,\delta_3,\delta_4\in\mathbb{R}}{\text{maximize}} & v(\delta_1,\delta_2,\delta_3,\delta_4) = \left(\dfrac{1}{\delta_1}\right)^{\delta_1}\left(\dfrac{1}{\delta_2}\right)^{\delta_2}\left(\dfrac{2}{\delta_3}\right)^{\delta_3}\left(\dfrac{1}{\delta_4}\right)^{\delta_4} & \\
\text{subject to} & 2\delta_1 \qquad + \delta_3 - \delta_4 = 0, & \text{(power of } t_1\text{)} \\
 & \qquad 2\delta_2 + \delta_3 - \delta_4 = 0, & \text{(power of } t_2\text{)} \\
 & \delta_1 + \ \ \delta_2 + \delta_3 + \delta_4 = 1, & \text{(AM-GM requirement)} \\
 & \delta_1, \delta_2, \delta_3, \delta_4 > 0. & \text{(AM-GM requirement)}
\end{array}
$$

The first two equations tell us $\delta_1 = \delta_2 = \frac{1}{2}(\delta_4 - \delta_3)$. So set $\delta_1 = \delta_2 = s$. Solving for $\delta_3$ and $\delta_4$ in terms of $s$ gives us a one-dimensional infinite family of solutions: $\boldsymbol{\delta} = (s, s, \frac{1}{2} - 2s, \frac{1}{2})$. All components of this vector must be positive, so we have $0 < s < \frac{1}{4}$.

Now comes the obnoxious step: maximizing $v(\boldsymbol{\delta})$. Since $\boldsymbol{\delta}$ is now a function of $s$, it is equivalent to maximize

$$
v(s) = \left(\frac{1}{s}\right)^s \left(\frac{1}{s}\right)^s \left(\frac{2}{\frac{1}{2} - 2s}\right)^{\frac{1}{2}-2s} \left(\frac{1}{\frac{1}{2}}\right)^{\frac{1}{2}}.
$$

It is easier (both in this case, and in general) to maximize $\log v(s)$ instead of $v(s)$, because the product becomes a sum:

$$
\begin{aligned}
\log v(s) &= s \log \frac{1}{s} + s \log \frac{1}{s} + \left(\frac{1}{2} - 2s\right) \log \frac{2}{\frac{1}{2} - 2s} + \frac{1}{2} \log \frac{1}{2} \\
&= -2s \log s + \left(2s - \frac{1}{2}\right) \log \left(\frac{1}{4} - s\right) - \frac{1}{2} \log 2.
\end{aligned}
$$

As usual, we find the critical point by setting the derivative to 0:

$$
\frac{d}{ds} \log v(s) = -2 \log s - 2 + 2 \log \left(\frac{1}{4} - s\right) + 2 = 0 \iff 2 \log \left(\frac{1}{4} - s\right) = 2 \log s.
$$

In general, this could get messy, but in this case, we can immediately conclude that $\frac{1}{4} - s = s$ and bypass dealing with the logarithms. This tells us that $s = \frac{1}{8}$ is the critical point.

In fact, the unique critical point will always be the global maximizer. The function $\log v(s) = \delta_1 \log \frac{C_1}{\delta_1} + \cdots + \delta_n \log \frac{C_n}{\delta_n}$ has Hessian matrix

$$
\begin{bmatrix}
-1/\delta_1 & 0 & \cdots & 0 \\
0 & -1/\delta_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & -1/\delta_n
\end{bmatrix}
$$

which is negative definite when $\delta_1, \ldots, \delta_n > 0$. (In other words, $\log v(\boldsymbol{\delta})$ is strictly concave.)

So the optimal dual solution is $\boldsymbol{\delta}^* = (\frac{1}{8}, \frac{1}{8}, \frac{1}{4}, \frac{1}{2})$, with $v(\boldsymbol{\delta}^*) = 8^{1/8} \cdot 8^{1/8} \cdot 8^{1/4} \cdot 2^{1/2} = 4$. This gives us the equations

$$
\frac{t_1^2}{1/8} = \frac{t_2^2}{1/8} = \frac{2t_1t_2}{1/4} = \frac{t_1^{-1}t_2^{-1}}{1/2} = 4
$$

giving the solution $\mathbf{t}^* = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$.