

Optimization Theory (DS2) Lecture #3

Three Kinds of Certificates, Standard Equality Form, and a Simplex Iteration

May 9, 2017

Abstract

Today we are going to prepare for the simplex algorithm. We will cover Secs. 2.1 to 2.3 of the textbook.

Text and examples adapted from A Gentle Introduction to Optimization.

1 Review

Recall, last time we worked on how to formulate problems, going from a word description of a problem to a formal description including **Variables**, **Constraints**, and the **Objective Function**. Once you have those, you can produce a formal description that says, “Maximize (some function) such that (s.t.) (some constraints hold).”

2 Geometric Interpretation of Constraints and Objective Function

First, let’s review the idea of constraints with a 2-D example. If I give you the equation

$$y \geq x + 2 \tag{1}$$

you know that this cuts the X-Y plane in half; points in one half satisfy the constraint, points in the other half don’t.

If we add a second constraint, ...

2.1 A Geometric Simplex

A *simplex* is an n -dimensional generalization of a triangle: in one dimension it’s a line segment, in two it’s a triangle, in three it’s a tetrahedron, in four or more it’s more generally called a simplex. It has $n + 1$ vertices. In terms of optimization problems, it has $n + 1$ independent constraints.

2.2 More Complex Polytopes

In the next lecture, we will see the n -dimensional generalization of a polyhedron known as a *polytope*.

3 Certificates

As the book says, “By definition, an LP [linear program] has only one optimal value, but it may have many optimal solutions.” However, it may not have a solution at all! Or, it may have a solution that runs off to infinity. So, there are three possible cases we would like to distinguish:

1. our problem is *infeasible* (has no solution),
2. it has one or more *optimal* solutions, or
3. it is *unbounded*.

Also, in the case where we have a solution that we think is optimal, it would be nice to be able to prove that to others. We will see that, once the problem is properly formulated, there are one or two vectors we can present that will show which of these cases holds. These vectors are called *certificates of infeasibility*, *optimality*, or *unboundedness*, respectively.

3.1 Infeasibility

We only need to look at the constraints, not the objective function, and demonstrate that they are mutually exclusive: by the time you take them all into account, there is no space left for a working solution at all.

3.2 2-D Example

To be added.

3.3 Larger Example

Consider the constraints of a problem (P),

$$4x_1 + 10x_2 - 6x_3 - 2x_4 = 6 \tag{2}$$

$$-2x_1 + 2x_2 - 4x_3 + x_4 = 5 \tag{3}$$

$$-7x_1 - 2x_2 + 4x_4 = 3, \tag{4}$$

$$x_1, x_2, x_3, x_4 \geq 0. \tag{5}$$

To show there is no solution, we create the equation

$$x_1 + 4x_2 + 2x_3 = -1 \tag{6}$$

and from this, we are done: all of the coefficients of the x_i are positive (or non-negative, including x_4), and all of the x_i are *constrained* to be non-negative, so there is no way this can be solved!

So how do we create that equation? Via the usual algebraic manipulation of the equations above:

$$1 \times (\text{Eq. 1}) - 2 \times (\text{Eq. 2}) + 1 \times (\text{Eq. 3}). \quad (7)$$

If we write down those values as $y_1 = 1$, $y_2 = -2$, $y_3 = 1$, then we can create the vector

$$\vec{y} = (y_1, y_2, y_3)^\top = (1, -2, 1)^\top. \quad (8)$$

The vector \vec{y} is said to be our *certificate of infeasibility*.

We can write this down as

$$A\vec{x} = \vec{b}, \quad (9)$$

where

$$A = \begin{pmatrix} 4 & 10 & -6 & -2 \\ -2 & 2 & -4 & 1 \\ -7 & -2 & 0 & 4 \end{pmatrix} \quad (10)$$

$$\vec{b} = \begin{pmatrix} 6 \\ 5 \\ 3 \end{pmatrix} \quad (11)$$

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}. \quad (12)$$

We multiply both sides by the vector \vec{y}^\top to get

$$\vec{y}^\top A\vec{x} = \vec{y}^\top \vec{b}, \quad (13)$$

Insert numbers for \vec{y} into that equation, and you will get an equation with all non-negative values on the left and a negative value on the right. Formally,

Let A be a matrix and \vec{b} be a vector. Then the system

$$A\vec{x} = \vec{b}, \vec{x} \geq \vec{0} \quad (14)$$

has no solution if there exists a vector \vec{y} s.t.

1. $\vec{y}^\top A \geq \vec{0}^\top$, and
2. $\vec{y}^\top \vec{b} < 0$.

Farkas' lemma tells us that there exists a certificate of infeasibility for every infeasible system.

3.4 Unbounded Linear Programs

To be added.

3.5 2-D Example

To be added.

3.6 Larger Example

Consider the already-formulated LP

$$\max\{z(\vec{x}) = \vec{c}^\top \vec{x} : A\vec{x} = \vec{b}, \vec{x} \geq \vec{0}\}, \quad (15)$$

where

$$A = \begin{pmatrix} 1 & 1 & -3 & 1 & 2 \\ 0 & 1 & -2 & 2 & -2 \\ -2 & -1 & 4 & 1 & 0 \end{pmatrix} \quad (16)$$

$$\vec{b} = \begin{pmatrix} 7 \\ -2 \\ -3 \end{pmatrix} \quad (17)$$

$$\vec{c} = \begin{pmatrix} -1 \\ 0 \\ 3 \\ 7 \\ -1 \end{pmatrix} \quad (18)$$

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}. \quad (19)$$

If we pick vectors $\vec{x}_a = (2, 0, 0, 1, 2)^\top$ and $\vec{d} = (1, 2, 1, 0, 0)^\top$ and define a function

$$\vec{x}(t) = \vec{x}_a + t\vec{d}, \quad (20)$$

then we can show that

$$\vec{c}^\top \vec{x}(t) = \vec{c}^\top \vec{x}_a + 2t, \quad (21)$$

and so as $t \rightarrow \infty$, so does our objective function. Formally,

Suppose there exists a problem (P) as above, with a feasible solution \vec{x}_a and a vector \vec{d} such that:

1. $A\vec{d} = \vec{0}$, and
2. $\vec{d} \geq \vec{0}$, and
3. $\vec{c}^\top \vec{d} > 0$,

then (P) is unbounded.

The pair of vectors \vec{d} and \vec{x}_a together are our *certificate of unboundedness*.

3.7 Optimality

See Sec. 2.1.3.

4 Standard Equality Form (SEF)

An LP is in *standard equality form* if:

1. it is a maximization problem;
2. other than the non-negativity constraints, all constraints are *equalities*; and
3. every variable has a non-negativity constraint.

That is, it can be written in the form

$$\max\{z(\vec{x}) = \vec{c}^T \vec{x} : A\vec{x} = \vec{b}, \vec{x} \geq \vec{0}\}. \quad (22)$$

Consider the LP

$$\max(1, -2, 4)(x_1, x_2, x_3)^T \quad (23)$$

subject to

$$\begin{pmatrix} 1 & 5 & 3 \\ 2 & -1 & 2 \\ 1 & 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \begin{matrix} \geq \\ \leq \\ = \end{matrix} \begin{pmatrix} 5 \\ 4 \\ 2 \end{pmatrix} \quad (24)$$

$$x_1, x_2 \geq 0. \quad (25)$$

x_3 is called a *free variable* since it doesn't have a non-negativity constraint. But a lot of our techniques and especially proofs depend on that. So, we introduce two new variables, x_3^+ and x_3^- , set $x_3 = x_3^+ - x_3^-$, and add constraints $x_3^+, x_3^- \geq 0$. After some algebra, you get

$$\max(1, -2, 4, -4)(x_1, x_2, x_3^+, x_3^-)^T \quad (26)$$

subject to

$$\begin{pmatrix} 1 & 5 & 3 & -3 \\ 2 & -1 & 2 & -2 \\ 1 & 2 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3^+ \\ x_3^- \end{pmatrix} \begin{matrix} \geq \\ \leq \\ = \end{matrix} \begin{pmatrix} 5 \\ 4 \\ 2 \end{pmatrix} \quad (27)$$

$$x_1, x_2, x_3^+, x_3^- \geq 0. \quad (28)$$

Getting closer, but we're still not quite there; we don't have equalities everywhere yet. So we introduce two *slack variables* x_4 and x_5 . Now we get

$$\max(1, -2, 4, -4, 0, 0)(x_1, x_2, x_3^+, x_3^-, x_4, x_5)^T \quad (29)$$

subject to

$$\begin{pmatrix} 1 & 5 & 3 & -3 & 0 & -1 \\ 2 & -1 & 2 & -2 & 1 & 0 \\ 1 & 2 & -1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3^+ \\ x_3^- \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \\ 2 \end{pmatrix} \quad (30)$$

$$x_1, x_2, x_3^+, x_3^-, x_4, x_5 \geq 0 \quad (31)$$

and we're finally in SEF! Note that when we replaced a \leq constraint (second line of Eq. 27), we added a 1 to the array, and when we replaced a \geq constraint (first line of Eq. 27), we added -1 to the array. This helps us keep our non-negativity constraint for all variables.

5 A Simplex Iteration

The idea of the *simplex algorithm* is pretty simple: repeat a *simplex operation* until no more simplex operations improve the result, and you're done. But there's a catch: the second simplex operation might partly undo the work of the first! So there is something you have to do in order to make that second one work, and that's what we'll see next time.

...Oh, wait, we haven't talked about a single simplex operation. Okay, let's talk about that. (n.b.: It's not *really* a simplex operation, in the mathematical sense, but it's related, and the name has stuck.)

Consider the following LP in SEF:

$$\max z(\vec{x}) = (2, 3, 0, 0, 0)(x_1, x_2, x_3, x_4, x_5)^\top \quad (32)$$

subject to

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 1 \end{pmatrix} \vec{x} = \begin{pmatrix} 6 \\ 10 \\ 4 \end{pmatrix} \quad (33)$$

$$x_1, x_2, x_3, x_4, x_5 \geq 0. \quad (34)$$

Given the solution $\vec{x}_a = (0, 0, 6, 10, 4)^\top$ (easy to see that's a solution – look at the right side of the array), $z(\vec{x}_a) = 0$.

Now try increasing x_1 , choosing $x_1 = t$. A little algebra gives

$$\begin{pmatrix} x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 6 \\ 10 \\ 4 \end{pmatrix} - t \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \quad (35)$$

from which we get

$$t \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \leq \begin{pmatrix} 6 \\ 10 \\ 4 \end{pmatrix} \quad (36)$$

which gives us the limiting inequality $t \leq 10/2 = 5$. Setting $t = 5$, $\vec{x}^t = (5, 0, 1, 0, 9)^\top$ and $z(\vec{x}^t) = 10$.

Unfortunately, we can't yet apply the same trick to x_2 , so that's the topic for next time!

6 Homework

See the separate file uploaded to SFS.

1. Exercise 1 (parts a, b) in Sec. 2.1.
2. Exercise 1 (part a) in Sec. 2.2.
3. Exercise 1 (all parts) in Sec. 2.3.