Math 482: Linear Programming[1] Mikhail Lavrov

Lecture 20: Fourier–Motzkin Elimination

March 13, 2020 *University of Illinois at Urbana-Champaign*

1 The feasibility problem

A related question to the linear program optimization problems we've been solving is the existence question: given A and $\mathbf{b}$, is the set $\{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \le \mathbf{b}\}$ nonempty? In other words, does there exist a *feasible* solution to a linear program?

We already have methods to solve this problem. Right now, one reasonable approach is to use the dual simplex method:

1. If necessary, put the constraints into equational form.
2. Find a basic solution (not necessarily a feasible one) by row reduction.
3. Invent an objective function that will give us a dual feasible tableau.
4. Use the dual simplex method to try to make the tableau primal feasible as well.

In fact, the feasibility problem is just as hard as the optimization problem. From strong duality, it follows that the optimization problem

$$\max\{\mathbf{c}^\mathsf{T}\mathbf{x} : A\mathbf{x} \le \mathbf{b}, \mathbf{x} \ge \mathbf{0}\}$$

can be rewritten as the problem of finding a point inside

$$\{(\mathbf{x}, \mathbf{u}) \in \mathbb{R}^{n+m} : \mathbf{c}^\mathsf{T}\mathbf{x} = \mathbf{u}^\mathsf{T}\mathbf{b}, A\mathbf{x} \le \mathbf{b}, \mathbf{u}^\mathsf{T}A \ge \mathbf{c}^\mathsf{T}, \mathbf{x} \ge \mathbf{0}, \mathbf{u} \ge \mathbf{0}\}.$$

So if we can solve any linear feasibility problem, we can solve any linear program as well.

2 Fourier–Motzkin elimination

There is another algorithm for solving the feasibility problem, which is very different from the simplex method. Here, "very different" is kind of like "worse", for large problems. But this algorithm, called Fourier–Motzkin elimination, is still occasionally interesting. It's simpler, so it's easier to prove things about. And for small problems, it might not even be that bad to use.

2.1 The algorithm

The idea of the algorithm is just this: first, we find a procedure to reduce an n-variable problem to an equivalent $(n-1)$-variable problem. (Equivalent means that one system of inequalities has a feasible solution if and only if the other one does.) Then, repeat this procedure, eliminating variables one at a time.

[1] This document comes from the Math 482 course webpage: https://faculty.math.illinois.edu/~mlavrov/courses/482-spring-2020.html

Eventually, we'll be left with a 1-variable problem, and these are easy to solve: just check if there are any numbers between the greatest lower bound and the least upper bound. We will even be able to trace back our steps, using a solution to the 1-variable problem to find solutions to the 2-variable, 3-variable, and eventually the original n-variable problem.

So how do we eliminate a variable? Well, let's suppose we have variables $x_1, x_2, \ldots, x_n$, and we want to eliminate x_n.

Our first step is to solve for x_n. For each inequality such as

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n \leq b$$

we will get one of the two inequalities

$$x_n \leq \frac{b - a_1x_1 - a_2x_2 - \cdots - a_{n-1}x_{n-1}}{a_n} \quad \text{or} \quad x_n \geq \frac{b - a_1x_1 - a_2x_2 - \cdots - a_{n-1}x_{n-1}}{a_n}$$

depending on whether $a_n > 0$ or $a_n < 0$. (If $a_n = 0$, we'll leave the inequality alone, since it doesn't involve x_n to begin with.)

We'll end up with a collection of lower and upper bounds:

$$x_n \geq L_1, x_n \geq L_2, \ldots, x_n \geq L_k \qquad x_n \leq U_1, x_n \leq U_2, \ldots, x_n \leq U_m.$$

Here, each L_i and U_j is an expression in the $n-1$ variables $x_1, x_2, \ldots, x_{n-1}$.

It's possible to choose a value of x_n that satisfies all these bounds if and only if

$$\max\{L_1, L_2, \ldots, L_k\} \leq \min\{U_1, U_2, \ldots, U_m\}$$

because thern there's something in between all the lower bounds and all the upper bounds. Unfortunately, we can't just write down this single equation, because the lower and upper bounds are all in terms of unknown values $x_1, \ldots, x_{n-1}$, so we can't tell which ones are larger or smaller. So we compromize by writing down $k \cdot m$ inequalities:

$$\begin{array}{cccc} L_1 \leq U_1 & L_1 \leq U_2 & \cdots & L_1 \leq U_m \\ L_2 \leq U_1 & L_2 \leq U_2 & \cdots & L_2 \leq U_m \\ \vdots & \vdots & \ddots & \vdots \\ L_k \leq U_1 & L_k \leq U_2 & \cdots & L_k \leq U_m \end{array}$$

If an x_n exists that satisfies all the constraints, then for every i and j, $L_i \leq x_m \leq U_j$, so all of these inequalities hold. Conversely, if all of these inequalities hold, then the max-min inequality holds as well, and we can choose x_n between $\max\{L_1, \ldots, L_k\}$ and $\min\{U_1, \ldots, U_m\}$.

This results in our new system in only $n-1$ variables. It consist of the $k \cdot m$ inequalities above, plus any of the original inequalities which didn't involve x_n in them to begin with.

2.2 Example

Suppose that we have a system of inequalities

$$\begin{cases} x - y \geq -3 \\ x + 2y \geq 4 \\ x + y \leq 7 \\ x, y \geq 0 \end{cases}$$

We begin by eliminating y. First, we solve for y in each inequality

$$\begin{cases} x - y \geq -3 \implies y \leq x + 3 \\ x + 2y \geq 4 \implies y \geq 2 - \frac{1}{2}x \\ x + y \leq 7 \implies y \leq 7 - x \\ y \geq 0 \\ x \geq 0 \end{cases}$$

Next, we pair up the two lower bounds on y with the two upper bounds. This gives us four inequalities on x, together with a fifth one which is just $x \geq 0$. These inequalities can all be simplified into lower or upper bounds on x:

$$\begin{cases} 2 - \frac{1}{2}x \leq 7 - x \\ 2 - \frac{1}{2}x \leq x + 3 \\ 0 \leq 7 - x \\ 0 \leq x + 3 \\ x \geq 0 \end{cases} \implies \begin{cases} x \leq 10 \\ x \geq -\frac{2}{3} \\ x \geq 7 \\ x \geq -3 \\ x \geq 0 \end{cases}$$

Taking the greatest lower bound and the least upper bound, we know that a solution exists for any $x \in [0, 7]$.

To find a value of y, we work backwards. Say we take $x = 2$. Then the inequalities on y become

$$\begin{cases} y \leq x + 3 \implies y \leq 5 \\ y \geq 2 - \frac{1}{2}x \implies y \geq 1 \\ y \leq 7 - x \implies y \leq 5 \\ y \geq 0 \end{cases}$$

So when $x = 2$, we can pick any $y \in [1, 5]$; for example, $y = 3$ gives the feasible solution $(2, 3)$.

2.3 Complexity

The downside of this method is that, in typical cases, the number of inequalities grows very quickly. For instance, here's one possible way that Fourier–Motzkin elimination can go:

$$\begin{aligned} 8 \text{ inequalities in } x_1, x_2, x_3, x_4 &\implies 4 \text{ lower and } 4 \text{ upper bounds on } x_4 \\ &\implies 4 \times 4 = 16 \text{ inequalities in } x_1, x_2, x_3 \end{aligned}$$

$$\begin{aligned}
&\implies 8 \text{ lower and } 8 \text{ upper bounds on } x_3 \\
&\implies 8 \times 8 = 64 \text{ inequalities in } x_1, x_2 \\
&\implies 32 \text{ lower and } 32 \text{ upper bounds on } x_2 \\
&\implies 32 \times 32 = 1024 \text{ inequalities in } x_1
\end{aligned}$$

The above is worst-case behavior: the types of bounds are split exactly in half every time. But we expect *roughly* half of each type of bound, so this is not too far off from typical behavior, either.

If we started with just 8 inequalities in n variables, then this pattern would continue to give us 2^{2^k+2} inequalities in $n-k$ variables all the way up until we get $2^{2^{n-1}+2}$ inequalities in 1 variable on the very last step. This is worse than exponential: this is doubly exponential.

3 Farkas's lemma

We can slightly modify the variable elimination procedure we used in Fourier–Motzkin elimination. Let's assume that all our inequalities are $\le$ inequalities, for consistency. Then, in the step where we want to eliminate the variable x_n:

- Instead of solving for x_n in each inequality, divide each inequality where x_n appears by $|a_n|$, the absolute value of x_n's coefficient. The result will have x_n either with a coefficient of 1 (corresponding to an upper bound) or -1 (corresponding to a lower bound).
- Instead of combining the lower and upper bounds, we add every $+x_n$ inequality to every $-x_n$ inequality. This gives us inequalities without x_n in them.

The inequalities we get this way will be the same ones, just in a different form. However, it is more straightforward to see that every inequality we get is a linear combination of the original inequalities.

Let's rewrite our previous example in this way, and annotate it so that we see how each inequality is obtained. Here is the "eliminate y" step:

$$\begin{array}{cc} (a) & -x+y\le 3 \\ (b) & -x-2y\le -4 \\ (c) & x+y\le 7 \\ (d) & -x\le 0 \\ (e) & -y\le 0 \end{array}
\rightsquigarrow
\begin{array}{cc} (a) & -x+y\le 3 \\ (c) & x+y\le 7 \\ \frac12(b) & -\frac12 x - y \le -2 \\ (e) & -y\le 0 \\ (d) & -x\le 0 \end{array}
\rightsquigarrow
\begin{array}{cc} (a)+\frac12(b) & -\frac32 x\le 1 \\ (a)+(e) & -x\le 3 \\ (c)+\frac12(b) & \frac12 x\le 5 \\ (c)+(e) & x\le 7 \\ (d) & -x\le 0 \end{array}$$

Then, we can eliminate x in the same way, getting many inequalities with no variables in them at all:

$$\begin{array}{cc} (a)+\frac12(b) & -\frac32 x\le 1 \\ (a)+(e) & -x\le 3 \\ \frac12(b)+(c) & \frac12 x\le 5 \\ (c)+(e) & x\le 7 \\ (d) & -x\le 0 \end{array}
\rightsquigarrow
\begin{array}{cc} (b)+2(c) & x\le 10 \\ (c)+(e) & x\le 7 \\ \frac23(a)+\frac13(b) & -x\le \frac23 \\ (a)+(e) & -x\le 3 \\ (d) & -x\le 0 \end{array}
\rightsquigarrow
\begin{array}{cc} \frac23(a)+\frac43(b)+2(c) & 0\le \frac{32}{3} \\ (a)+(b)+2(c)+(e) & 0\le 13 \\ (b)+2(c)+(d) & 0\le 10 \\ \frac23(a)+\frac13(b)+(c)+(e) & 0\le \frac{23}{3} \\ (a)+(c)+2(e) & 0\le 10 \\ (c)+(d)+(e) & 0\le 7 \end{array}$$

In this particular example, we gain nothing new from doing this.

However, imagine, that we started with a different system of equations which actually had no feasible solution. For example, change equation (c) from $x+y \leq 7$ to $x+y \leq 1$. Then we'd still get the combination $(b)+2(c)+(d)$ at the end, but instead of saying $0 \leq 10$, it would say $0 \leq -2$. In other words, we'd have deduced a contradiction as a combination of our starting inequalities.

This is very convenient for obtaining a short proof that the system of equations has no feasible solution, rather than having to say something like "we tried to find a solution and we couldn't".

The theoretical principle that any system of inequalities either has a solution or a short proof of non-existence is known as Farkas's lemma:

Theorem 3.1 (Farkas's lemma)**.** *For any system of inequalities $A\mathbf{x} \leq \mathbf{b}$, either there exists a feasible solution $\mathbf{x}$, or we can take a linear combination of the inequalities to derive a contradiction: there exists a vector $\mathbf{u} \geq \mathbf{0}$ such that $\mathbf{u}^\mathsf{T} A = \mathbf{0}^\mathsf{T}$ but $\mathbf{u}^\mathsf{T}\mathbf{b} < 0$.*

(The existence of $\mathbf{u}$ is the same as being able to derive a contradiction: left-multiplying $A\mathbf{x} \leq \mathbf{b}$ by $\mathbf{u}^\mathsf{T}$, we deduce $\mathbf{0}^\mathsf{T}\mathbf{x} < 0$, or $0 < 0$.)

Proof. Fourier–Motzkin elimination always gives us either a feasible solution $\mathbf{x}$ or the vector $\mathbf{u}$. □

Farkas's lemma is important because it is the equivalent of strong duality for the feasibility problem: the system "$\mathbf{u}^\mathsf{T} A = \mathbf{0}^\mathsf{T}$ and $\mathbf{u}^\mathsf{T}\mathbf{b} < 0$ and $\mathbf{u} \geq \mathbf{0}$" is a dual feasibility problem to the primal feasibility problem "$A\mathbf{x} \leq \mathbf{b}$". We could have used Farkas's lemma to prove strong duality for linear programs, without relying on the correctness of the simplex method to get there.