451: Linear Programming II

G. Miller, K. Sutner
Carnegie Mellon University
2020/11/10

A linear program in general form has

Objective minimize or maximize

Constraints $\leq$, $=$, $\geq$

Variables non-negative, non-positive, unconstrained

We will now introduce a notion of duality that associates a minimization LP with a corresponding maximization LP. For general form this is a bit messy, see below, so let us focus on an LP of the form

$$\begin{aligned} \max\ c \circ x \\ Ax \leq b \\ x \geq 0 \end{aligned}$$

Computing the exact solution of an LP can be expensive, so it is natural to look for lower/upper bounds on the value. How would one go about doing this?

For the last LP, any feasible point x automatically produces a lower bound on the optimal value.

How do we get an upper bound: $c \circ x \leq \beta$ for all feasible x?

Here is another Ansatz: suppose our objective function is $5x_1 + 18x_2$ and there is a constraint

$$x_1 + 3x_2 \leq 7$$

Multiplying the constraint by 6 and exploiting non-negativity we get

$$5x_1 + 18x_2 \quad \leq \quad 6x_1 + 18x_2 \quad \leq \quad 42$$

Note that multiplier 6 is optimal because of x_2.

More generally, we could take any linear combinations of the rows of $(A \mid b)$. Since $A_i^{\rightarrow} \circ x \leq y_i$ we can choose $y \in \mathbb{R}^m$ as coefficients of a linear combination:

$$y^T A x = \sum y_i (A_i^{\rightarrow} \circ x) \leq \sum y_i b_i = y \circ b$$

The goal now is to choose y in such a way that the RHS becomes as small as possible (while staying above the objective function).

Key Idea: The question of finding the "optimal" y for this purpose can itself be expressed as a linear program!

Primal Problem (P)

$$\begin{aligned} \text{max } c \circ x \\ Ax \leq b \\ x \geq 0 \end{aligned}$$

Dual Problem (D)

$$\begin{aligned} \text{min } y \circ b \\ A^T y \geq c \\ y \geq 0 \end{aligned}$$

Here $A \in \mathbb{R}^{m \times n}$, $m \leq n$, $y, b \in \mathbb{R}^m$, $x, c \in \mathbb{R}^n$.

Exercise

Check in detail that this really corresponds to the upper bound motivation outlined above.

Suppose we want to dualize a general form maximization problem into a general form minimization problem. The recipe is this:

	primal	dual	
objective	$\max c \circ x$	$\min y \circ b$	
	$a \circ x \leq b_i$	$y_i \geq 0$	
constraints	$a \circ x = b_i$	y free	variables
	$a \circ x \geq b_i$	$y_i \leq 0$	
	$x_j \geq 0$	$a^T \circ y \geq c_j$	
variables	x free	$a^T \circ y = c_j$	constraints
	$x_j \leq 0$	$a^T \circ y \geq c_j$	

Consider an equality constraint

$$a \circ x = b_i$$

We can translate this into two inequalities

$$\begin{aligned} a \circ x &\leq b_i \\ -a \circ x &\leq -b_i \end{aligned}$$

which then turn into two dual variables y_i^+ and y_i^- and a term $b_i(y_i^+ - y_i^-)$ in the dual objective function.

But then we could equivalently just use one free dual variable.

Lemma

The dual of the dual is the primal.

Exercise

How about dualizing a minimization problem?

Exercise

Prove the dualization lemma.

Exercise

Consider the three possibilities: optimal solution, unbounded, infeasible.

What combinations of the these are possible between the primal and dual?

Theorem ($\min = \max$)

If the primal and dual both have a solution, then their values agree.

Goes back to von Neumann, Dantzig 1947 and Gale, Kuhn, Tucker 1951.

Generalizes for example the max-flow-min-cut theorem.

The direction $\max \leq \min$ (weak duality) is easy:

$$c \circ x \leq (A^T y) \circ x = y^T A\, x = y^T b = y \circ b$$

Alas, the opposite direction requires a bit more work.

We need a little background from analysis.

Theorem (Bolzano-Weierstrass 1817)

A continuous function on a compact domain assumes its minimum.

Note that this provides another proof for the existence of an optimal solution in the case where the LP is feasible and bounded: the simplex is a compact set.

Our vertex-hopping proof from last time is much more constructive and provides an actual starting point for algorithmic purposes.

Lemma

Let $\emptyset \neq X \subseteq \mathbb{R}^n$ be closed and convex, $z \notin X$. Then there exists a point $\widehat{x} \in X$ that minimizes the distance $\|z - x\|$.
Moreover, $(z - \widehat{x}) \circ (x - \widehat{x}) \leq 0$ for all $x \in X$.

Proof.

By Bolzano-Weierstrass, the distance function $\|z - x\|$ assumes its minimum at some point $\widehat{x} \in X$.

Since X is convex we have $\widehat{x} + \lambda(x - \widehat{x}) \in X$ for all $\lambda \in [0, 1]$. But then

$$\begin{aligned} \|z - \widehat{x}\|^2 &\leq \|(z - \widehat{x}) - \lambda(x - \widehat{x})\|^2 \\ &= \|z - \widehat{x}\|^2 + \lambda^2 \|x - \widehat{x}\|^2 - 2\lambda(z - \widehat{x}) \circ (x - \widehat{x}) \end{aligned}$$

Hence $2(z - \widehat{x}) \circ (x - \widehat{x}) \leq \lambda \|x - \widehat{x}\|^2$ and letting $\lambda \to 0$ we get our claim.

□

Corollary

Let $\emptyset \neq X \subseteq \mathbb{R}^n$ be closed and convex, $z \notin X$. Then there exists a hyperplane H that separates z and X.

Proof.

Recall $(z - \widehat{x}) \circ (x - \widehat{x}) \leq 0$ for all $x \in X$ (obtuse angle $\measuredangle z\widehat{x}x$).

Define $u = \widehat{x} - z$ and $\mu = u \circ \widehat{x}$, so $u \circ x \geq \mu$.

Then $H = \{ x \in \mathbb{R}^n \mid u \circ x = \mu \}$ works.

□

Note that $X \subseteq H^+ = \{ x \in \mathbb{R}^n \mid u \circ x \geq \mu \}$ and $z \in H^-$.

Lemma

Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$. *Then either*

1. $A\,x = b$, $x \geq 0$ *for some* x, *or*
2. $A^T y \geq 0$, $y \circ b < 0$ *for some* y.

Proof. Clearly the two conditions contradict each other:
$0 > y \circ b = y^T A\,x \geq 0$.

Now assume that (1) fails and set $X = \{\, A\,x \mid x \geq 0 \,\}$. Then $b \notin X$, and X is closed and convex. So there must be a separating hyperplane $H = \{\, z \in \mathbb{R}^n \mid u \circ z = \mu \,\}$: $u \circ b < \mu$ and $u \circ z \geq \mu$ for all $z \in X$.

Hence $u^T A\,x \geq \mu$ for all $x \geq 0$. But x is unbounded, so $u^T A \geq 0$.

Done letting $y = u$. □

For our application to duality, it is convenient to use the following simple corollary to Farkas' lemma.

- Either there exists an x such that

$$\begin{aligned} A\,x &\leq b \\ x &\geq 0 \end{aligned}$$

- or there exists a y such that

$$\begin{aligned} A^T y &\geq 0 \\ y \circ b &< 0 \\ y &\geq 0 \end{aligned}$$

Just add slack variables to the inequalities in the original lemma.

We still have to show that $\min \leq \max$. It suffices to show $\max < \beta$ implies $\min < \beta$.

We express the condition $\max < \beta$ as an LP and consider it and its dual:

$$
\begin{aligned}
A\,x &\leq b \\
-c \circ x &\leq -\beta \\
x &\geq 0
\end{aligned}
\qquad\qquad
\begin{aligned}
A^T y - cv &\geq 0 \\
y \circ b - \beta v &< 0 \\
y, v &\geq 0
\end{aligned}
$$

By our assumption, the first system is not feasible.

By the corollary to Farkas, the second systems has a solution $y, v \geq 0$.

Case 1: $v = 0$

Then the following is feasible:

$$\begin{aligned} A^T y &\geq 0 \\ y \circ b &< 0 \\ y &\geq 0 \end{aligned}$$

By Farkas, the original primal has no solution, contradiction.

Case 2: $v > 0$

By scaling we can get $v = 1$.

But then y solves the original dual, and $y \circ b < \beta$.

$\square$

An alternative proof of strong duality can be based on a careful analysis of the state of the simplex algorithm, and in particular the last state after termination.

Alas, the arithmetic is somewhat messy and rather tedious. As always, theory helps.

Note that the only obvious upper bound on the number of rounds is $\binom{n}{m}$.

Theorem (Klee, Minty 1972)

There are instances of LP with exponentially many rounds.

Careful, this depends on implementation details (choice of pivot). Also, the slow examples tend to be rather artificial.

Theorem (Smale 1983)

The average running time of Simplex is polynomial.

There is a question whether the distribution used by Smale is relevant in practice. It seems that often something like $2(n + m)$ rounds suffice.

To make sense out of the complexity of LP we have two choices:

Rationalize Use rational numbers instead of reals. Let L be the total number of bits needed for these rationals, then worry about standard complexity in terms of n, m and L.
For example, we could ask if the problem is in $\mathbb{P}$. It is.

Real Computation Switch to a model of computation that accommodates the reals (computability over algebraic structures). This may sound crazy, but e.g. the first-order theory of the reals is decidable (Tarski), but for integers it is undecidable.

We will ignore the second option.

We may as well assume that the input is given in terms of integers: So an instance looks like

$$A \in \mathbb{Z}^{m \times n}, b \in \mathbb{Z}^m, c \in \mathbb{Z}^n$$

We need to

- report "infeasible" if the simplex is empty,
- report "unbounded" if the objective function is unbounded on the simplex,
- return an optimal solution, otherwise.

One can also concoct decision problems: "is the LP feasible?" or "is there a feasible x of value at least/most β".

It is tempting to assume that the rational version of LP is in $\mathbb{NP}$: we could guess a solution and then verify that it satisfies the given bound.

But there is a problem: the numbers must not get too large, we cannot use an exponential number of bits.

We are saved by the fact that we are dealing with linear algebra here: one can check that the total number of bits needed is polynomial.

Exercise

Use Cramer's Rule to show that this is indeed the case.

Theorem

Linear programming is in $\mathbb{NP} \cap$ *co-*$\mathbb{NP}$*.*

For the proof one can exploit Farkas' lemma to obtain a certificate for a no-instance.

Heuristically, once a problem drops into $\mathbb{NP} \cap$ co-$\mathbb{NP}$, it is often on its way down to $\mathbb{P}$. A perfect example is primality (disregarding efficiency issues).

Exercise

Figure out the details.

Theorem (Khachian 1979)

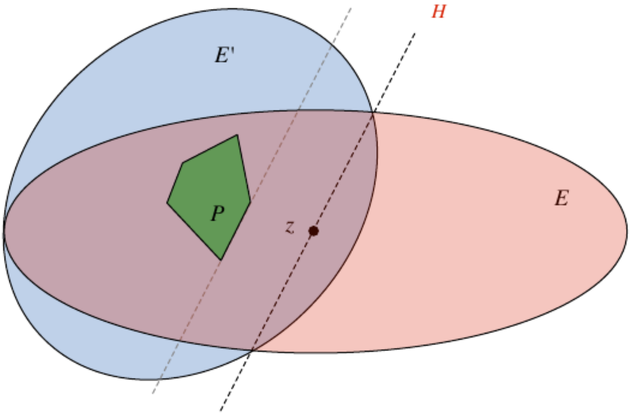
Linear programming can be solved in polynomial time.

Alas, this result had no direct impact on practical algorithms.

The technique used is the so-called ellipsoid method (an ellipsoid is an affine transform of the unit sphere). The idea is to find a point inside the simplex, and then use that point to find a vertex of optimal value.

To find the interior point, enclose the simplex in an ellipsoid E. If the centroid of E lies in the simplex, stop. Otherwise, replace E by a smaller ellipsoid E'.

Turns out to be $O(n^6 L^2)$.


H
E'
P
z
E

Khachian's algorithm works with a sequence of exterior points to find an interior one. By contrast, Karmarkar uses a sequence of interior points that converges to optimality (and then finds an optimal vertex).

Uses a particular normal form, a specialized standard form.

- $b_i = 0$ for $i < m$, $b_m = 1$
- $a_{mj} = 1$.
- $(1/n)_n$ is feasible, every feasible point has non-negative value.

The goal is to find a feasible point of value 0. Other forms can be transformed to this normal form in liner time.

The algorithm turns out to be practical, with running time $O(n^{3.5}L^2)$.

It even lead to a patent fight (thank you, AT&T), and specialized hardware.

It is tempting to ask what happens if we try to solve a Linear Program when the variables are required to range over $\mathbb{Z}$.

As Matiyasevic has shown, solving multivariate polynomial equations over $\mathbb{Z}$ turns out to be hugely more difficult than over $\mathbb{R}$: a (highly nontrivial) decidable problem goes rogue and becomes undecidable.

However, we are saved by the fact that we are dealing with linear algebra here. As before in the rational case, the problem is in $\mathbb{NP}$: the number of bits needed to represent the relevant integers stays polynomially bounded.

Variables: x_v indicator variable for each Boolean variable v

Constraints: $0 \leq x_v \leq 1$, $1 \leq x' + y' + z'$ for each clause $\{x, y, z\}$

Minimize: $\sum x_v$.

Here $x' = x$ if x appears positively, $x' = 1 - x$ otherwise. To get a decision problem, distinguish between feasible and not feasible.

This is an example of 0/1-Integer Programming: the variables are constrained to $\mathbf{2}$ (and membership in $\mathbb{NP}$ is trivial).

Claim

0/1-Integer Programming is $\mathbb{NP}$-complete.