6.854J / 18.415J Advanced Algorithms
Fall 2008

9 Linear Programming

Linear programming is the class of optimization problems consisting of optimizing the value of a linear objective function, subject to linear equality or inequality constraints. These constraints are of the form

$$a_1 x_1 + \cdots + a_n x_n \quad \{\leq, =, \geq\} \quad b,$$

where $a_i, b \in \mathbb{R}$, and the goal is to maximize or minimize an objective function of the form

$$c_1 x_1 + \cdots + c_n x_n.$$

In addition, we constrain the variables x_i to be nonnegative.

The problem can be expressed in matrix form. Given these constraints

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

maximize or minimize the value of

$$\mathbf{c}^T \mathbf{x},$$

where $\mathbf{x} \in \mathbb{R}^n, \mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m, \mathbf{c} \in \mathbb{R}^n$.

Linear programming has many applications and can also be used as a proof technique. In addition, it is important from a complexity point-of-view, since it is among the hardest of the class of polynomial-time solvable problems.

9.1 Algorithms

Research in linear programming algorithms has been an active area for over 60 years. In this class, we will discuss three major (classes of) algorithms:

- Simplex method (Dantzig 1947).

  - Fast in practice.
  - Still the most-used LP algorithm today.
  - Can be nonpolynomial (exponential) in the worst case.

- Ellipsoid algorithm (Shor, Khachian 1979).

  - Polynomial time; this was the first polynomial-time algorithm for linear programming.
  - Can solve LP (and other more general) problems where the feasible region $P = \{x : Ax = b, x \geq 0\}$ is not explicitly given, but instead, given a vector x, one can efficiently decide whether $x \in P$ or if not, find an inequality satisfied by P but not by x.
  - Very useful for designing polynomial time algorithms for other problems.
  - Not fast in practice.

- Interior-point algorithms (Karmarkar 1984).

  - This is a class of algorithms which maintain a feasible point in the interior of P; many variants (by many researchers) have been developed.
  - Polynomial time.
  - Fast in practice.
  - Can beat the simplex method for larger problems.

9.2 Equivalent forms

A linear programming problem can be modified to fit a preferred alternate form by changing the objective function and/or the linear constraints. For example, one can easily transform any linear program into teh standard form: $\min\{c^T x : Ax = b, x \geq 0\}$. One can use the following simple transformations.

Maximize to minimize $\qquad \max\{\mathbf{c}^T\mathbf{x}\} \longrightarrow \min\{-\mathbf{c}^T\mathbf{x}\}$

Equality to inequality $\qquad \mathbf{a}_i^T\mathbf{x} = b_i \longrightarrow \begin{cases} \mathbf{a}_i^T\mathbf{x} \leq b_i \\ \mathbf{a}_i^T\mathbf{x} \geq b_i \end{cases}$

Inequality to nonnegativity constraint $\qquad \mathbf{a}_i^T\mathbf{x} \leq b_i \longrightarrow \begin{cases} \mathbf{a}_i^T\mathbf{x} + s = b_i \\ s \geq 0 \end{cases} \quad (s \in \mathbb{R}^n)$

Variables unrestricted in sign $\qquad x_j \text{ unrestricted in sign} \longrightarrow \begin{cases} \text{replace } x_j \text{ everywhere by } x_j^+ - x_j^- \\ x_j^+ \geq 0 \\ x_j^- \geq 0 \end{cases}$

9.3 Definitions

Here is some basic terminology for a linear program.

Definition 1 *A vector x is* feasible *for an LP if it satisfies all the constraints.*

Definition 2 *An LP is* feasible *if there exists a feasible solution x for it.*

Definition 3 *An LP is* infeasible *if there is no feasible solution x for it.*

Definition 4 *An LP $\min\{\mathbf{c}^T\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0\}$ is* unbounded *if, for all $\lambda \in \mathbb{R}$, $\exists \mathbf{x} \in \mathbb{R}^n$ such that*

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$
$$\mathbf{x} \geq 0$$
$$\mathbf{c}^T\mathbf{x} \leq \lambda.$$

9.4 Farkas' lemma

If we have a system of equations $\mathbf{A}\mathbf{x} = \mathbf{b}$, from linear algebra, we know that either $\mathbf{A}\mathbf{x} = \mathbf{b}$ is solvable, or the system $\mathbf{A}^T\mathbf{y} = 0$, $\mathbf{b}^T y \neq 0$ is solvable. Indeed, since $Im(\mathbf{A}) = ker(\mathbf{A}^T)^\perp$, either $\mathbf{b}$ is orthogonal to $ker(\mathbf{A}^T)$ (in which case it is in the image of $\mathbf{A}$, i.e. $\mathbf{A}\mathbf{x} = \mathbf{b}$ is solvable) or it is not orthogonal to it in which case one can find a vector $\mathbf{y} \in \ker(\mathbf{A}^T)$ with a non-zero inner product with $\mathbf{b}$ (i.e. $\mathbf{A}^T\mathbf{y} = 0$, $\mathbf{b}^T y \neq 0$ is solvable).

Farkas' lemma generalizes this when we have also linear inequalities:

Lemma 1 ((Farkas' lemma)) *Exactly one of the following holds:*

1. $\exists \mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0$,

2. $\exists \mathbf{y} \in \mathbb{R}^m : \mathbf{A}^T \mathbf{y} \geq 0, \mathbf{b}^T \mathbf{y} < 0$.

Clearly, both cannot simultaneously happen, since the existence of such an $\mathbf{x}$ and a such a $\mathbf{y}$ would mean:

$$\mathbf{y}^T \mathbf{A} \mathbf{x} = \mathbf{y}^T (\mathbf{A} \mathbf{x}) = y^T \mathbf{b} < 0,$$

while

$$\mathbf{y}^T \mathbf{A} \mathbf{x} = (\mathbf{A}^T \mathbf{y})^T \mathbf{x} \geq 0,$$

as the inner product of two nonnegative vectors is nonnegative. Together this gives a contradiction.

9.4.1 Generalizing Farkas' Lemma

Before we provide a proof of the (other part of) Farkas' lemma, we would like to briefly mention other possible generalizations of the solvability of system of equations.

First of all, consider the case in which we would like the variables $\mathbf{x}$ to take integer values, but don't care whether they are nonnegative or not. In this case, the natural condition indeed is necessary and sufficient. Formally, suppose we take this set of constraints:

$$\begin{aligned} \mathbf{A}\mathbf{x} &= \mathbf{b} \\ \mathbf{x} &\in \mathbb{Z}^n \end{aligned}$$

Then if $\mathbf{y}^T \mathbf{A} \mathbf{x} = \mathbf{y}^T \mathbf{b}$, and we can find some $\mathbf{y}^T \mathbf{A} \in \mathbb{Z}^n$ and some $\mathbf{y}^T \mathbf{b}$ that is not integral, then the system of constraints is infeasible. The converse is also true.

Theorem 2 *Exactly one of the following holds:*

1. $\exists \mathbf{x} \in \mathbb{Z}^n : \mathbf{A}\mathbf{x} = \mathbf{b}$,

2. $\exists \mathbf{y} \in \mathbb{R}^m : \mathbf{A}^T \mathbf{y} \in \mathbb{Z}^n$ and $\mathbf{b}^T \mathbf{y} \notin \mathbb{Z}$.

One could try to combine both nonnegativity constraints and integral restrictions but in that case, the necessary condition for feasibility is not sufficient. In fact, for the following set of constraints:

$$\begin{aligned} \mathbf{A}\mathbf{x} &= \mathbf{b} \\ \mathbf{x} &\geq 0 \\ \mathbf{x} &\in \mathbb{Z}^n, \end{aligned}$$

determining feasibility is an NP-hard problem, and therefore we cannot expect a *good characterization* (a necessary and sufficient condition that can be checked efficiently).

9.4.2 Proof of Farkas' lemma

We first examine the projection theorem, which will be used in proving Farkas' lemma (see Figure 1).

Theorem 3 (The projection theorem) *If K is a nonempty, closed, convex set in $\mathbb{R}^m$ and $\mathbf{b} \notin K$, define*

$$\mathbf{p} = \operatorname{proj}_K(\mathbf{b}) = \arg \min_{\mathbf{z} \in K} \|\mathbf{z} - b\|_2. \tag{1}$$

Then, for all $\mathbf{z} \in K : (\mathbf{z} - \mathbf{p})^T (\mathbf{b} - \mathbf{p}) \leq 0$.
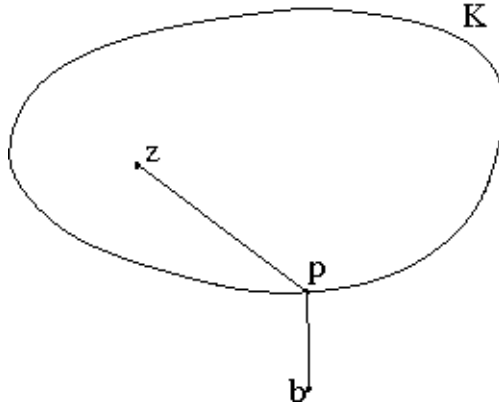
Figure 1: The projection theorem.

Proof of Lemma 1:   We have seen that both systems cannot be simultaneously solvable.

So, now assume that $\nexists \mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0$ and we would like to show the existence of $\mathbf{y}$ satisfying the required conditions. Define

$$K = \{\mathbf{A}\mathbf{x} : \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \geq 0\} \subseteq \mathbb{R}^m.$$

By assumption, $\mathbf{b} \notin K$, and we can apply the projection theorem. Define $\mathbf{p} = \mathrm{proj}_K(\mathbf{b})$. Since $\mathbf{p} \in K$, we have that $\mathbf{p} = \mathbf{A}\mathbf{x}$ for some vector $\mathbf{x} \geq 0$. Let $\mathbf{y} = \mathbf{p} - \mathbf{b} \in \mathbb{R}^m$. We claim that $\mathbf{y}$ satisfies the right conditions.

Indeed, consider any point $\mathbf{z} \in K$. We know that $\exists \mathbf{w} \geq 0 : \mathbf{z} = \mathbf{A}\mathbf{w}$. By the projection theorem, we have that $(\mathbf{A}\mathbf{w} - \mathbf{A}\mathbf{x})^T \mathbf{y} \geq 0$, i.e.

$$(\mathbf{w} - \mathbf{x})^T \mathbf{A}^T \mathbf{y} \geq 0, \tag{2}$$

for all $\mathbf{w} \geq 0$. Choosing $\mathbf{w} = \mathbf{x} + e_i$ (where e_i is the ith unit vector), we see that $\mathbf{A}^T \mathbf{y} \geq 0$. We still need to show that $\mathbf{b}^T y < 0$. Observe that $\mathbf{b}^T \mathbf{y} = (\mathbf{p} - \mathbf{y})^T \mathbf{y} = \mathbf{p}^T \mathbf{y} - \mathbf{y}^T \mathbf{y} < 0$ because $\mathbf{p}^T \mathbf{y} \leq 0$ and $\mathbf{y}^T \mathbf{y} > 0$. The latter follows from $y \neq 0$ and the former from (2) with $\mathbf{w} = 0$: $-\mathbf{x}^T \mathbf{A}^T \mathbf{y} \geq 0$, i.e. $-\mathbf{p}^T \mathbf{y} \geq 0$. $\qquad\qquad\qquad\square$

9.4.3   Corollary to Farkas' lemma

Farkas' lemma can also be written in other equivalent forms.

Corollary 4 *Exactly one of the following holds:*

1. $\exists \mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \leq \mathbf{b}$,

2. $\exists \mathbf{y} \in \mathbb{R}^m : \mathbf{y} \geq 0, \mathbf{A}^T y = 0, \mathbf{b}^T y < 0$.

Again, $\mathbf{x}$ and $\mathbf{y}$ cannot simultaneously exist. This corollary can be either obtained by massaging Farkas' lemma (to put the system of inequalities in the right form), or directly from the projection theorem.

9.5   Duality

Duality is one of the key concepts in linear programming. Given a solution $\mathbf{x}$ to an LP of value z, how do we decide whether or not $\mathbf{x}$ is in fact an optimum solution? In other words, how can we calculate a lower bound on $\min \mathbf{c}^T \mathbf{x}$ given that $\mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0$?

Suppose we have $\mathbf{y}$ such that $\mathbf{A}^T\mathbf{y} \leq \mathbf{c}$. Then observe that $\mathbf{y}^T\mathbf{b} = \mathbf{y}^T\mathbf{A}\mathbf{x} \leq \mathbf{c}^T\mathbf{x}$ for any feasible solution $\mathbf{x}$. Thus $\mathbf{y}^T\mathbf{b}$ provides a lower bound on the value of our linear program. This conclusion is true for all $\mathbf{y}$ satisfying $\mathbf{A}^T\mathbf{y} \leq \mathbf{c}$, so in order to find the best lower bound, we wish to maximize $\mathbf{y}^T\mathbf{b}$ under the constraint of $\mathbf{A}^T\mathbf{y} \leq \mathbf{c}$.

We can see that this is in fact itself another LP. This new LP is called the *dual* linear program of the original problem, which is called the *primal* LP.

- Primal LP: $\min \mathbf{c}^T\mathbf{x}$, given $\mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0$,

- Dual LP: $\max \mathbf{b}^T\mathbf{y}$, given $\mathbf{A}^T\mathbf{y} \leq \mathbf{c}$.

9.5.1   Weak Duality

The argument we have just given shows what is known as weak duality.

Theorem 5 *If the primal P is a minimization linear program with optimum value $\mathbf{z}$, then it has a dual D, which is a maximization problem with optimum value $\mathbf{w}$ and $\mathbf{z} \geq \mathbf{w}$.*

Notice that this is true even if either the primal or the dual is infeasible or unbounded, provided we use the following convention:

$$
\begin{array}{rcl}
\text{infeasible min. problem} & \longrightarrow & value = +\infty \\
\text{unbounded min. problem} & \longrightarrow & value = -\infty \\
\text{infeasible max. problem} & \longrightarrow & value = -\infty \\
\text{unbounded max. problem} & \longrightarrow & value = +\infty
\end{array}
$$

9.5.2   Strong Duality

What is remarkable is that one even has strong duality, namely both linear programs have the same values provided at least one of them is feasible (it can happen that both the primal and the dual are infeasible).

Theorem 6 *If P or D is feasible, then $\mathbf{z} = \mathbf{w}$.*

Proof: We assume that P is feasible (the argument if D is feasible is analogous; or one could also argue that the dual of the dual is the primal and therefore one can exchange the roles of primal and dual).

If P is unbounded, $\mathbf{z} = -\infty$, and by weak duality, $\mathbf{w} \leq \mathbf{z}$. So it must be that $\mathbf{w} = -\infty$ and thus $\mathbf{z} = \mathbf{w}$.

Otherwise (if P is not unbounded), let $\mathbf{x}^*$ be the optimum solution to P, i.e.:

$$
\begin{array}{rcl}
\mathbf{z} & = & \mathbf{c}^T\mathbf{x}^* \\
\mathbf{A}\mathbf{x}^* & = & \mathbf{b} \\
\mathbf{x}^* & \geq & 0
\end{array}
$$

We would like to find a dual feasible solution with the same value as (or no worse than) $\mathbf{x}^*$. That is, we are looking for a $\mathbf{y}$ satisfying:

$$
\begin{array}{rcl}
\mathbf{A}^T\mathbf{y} & \leq & \mathbf{c} \\
\mathbf{b}^T\mathbf{y} & \geq & \mathbf{z}
\end{array}
$$

If no such $\mathbf{y}$ exists, we can use Farkas' lemma to derive: $\exists \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \geq 0$, and $\exists \lambda \in \mathbb{R}, \lambda \geq 0$: $\mathbf{A}\mathbf{x} - \lambda\mathbf{b} = 0$ and $\mathbf{c}^T\mathbf{x} - \lambda\mathbf{z} < 0$.

We now consider two cases.

- If $\lambda \neq 0$, we can scale by λ, and therefore assume that $\lambda = 1$. Then we get that

$$\exists \mathbf{x} \in \mathbb{R}^n : \begin{cases} \mathbf{A}\mathbf{x} = \mathbf{b}, \\ \mathbf{x} \geq 0 \\ \mathbf{c}^T \mathbf{x} < \mathbf{z}. \end{cases}$$

This result is a contradiction, because $\mathbf{x}^*$ was the optimum solution, and therefore we should not be able to further minimize $\mathbf{z}$.

- If $\lambda = 0$ then

$$\exists \mathbf{x} \in \mathbb{R}^m : \begin{cases} \mathbf{x} \geq 0 \\ \mathbf{A}\mathbf{x} = 0 \\ \mathbf{c}^T \mathbf{x} < 0. \end{cases}$$

Consider now $\mathbf{x}^* + \mu\mathbf{x}$ for any $\mu > 0$. We have that

$$\begin{aligned} \mathbf{x}^* + \mu\mathbf{x} &\geq 0 \\ \mathbf{A}(\mathbf{x}^* + \mu\mathbf{x}) &= \mathbf{A}\mathbf{x}^* + \mu\mathbf{A}\mathbf{x} = \mathbf{b} + 0 = \mathbf{b}. \end{aligned}$$

Thus, $\mathbf{x}^* + \mu\mathbf{x}$ is feasible for any $\mu \geq 0$. But, we have that

$$\mathbf{c}^T(\mathbf{x}^* + \mu\mathbf{x}) = \mathbf{c}^T\mathbf{x}^* + \mu\mathbf{c}^T\mathbf{x} < z,$$

a contradiction.

$\square$