Analysis of Algorithms	Strong LP Duality
CS6820 Fall 2022	Monday, December 5, 2022

In this lecture we show that any LP maximization problem and its dual

$$\begin{array}{ll} \text{maximize } c^T x & \text{minimize } y^T b \\ \text{subject to } Ax \leq b,\ x \geq 0 & \text{subject to } y^T A \geq c^T,\ y \geq 0 \end{array} \tag{1}$$

have the same optimal value.

1 Slack Variables

The LP maximization problem and its dual above can be converted to an equivalent LP with equality constraints by adding *slack variables*. Suppose $A \in \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^n$, and $b \in \mathbb{R}^m$. For each constraint $a_{i1}x_1 + \cdots + a_{in}x_n \leq b_i$ corresponding to the ith row of A, we add a slack variable w_i and replace the constraint with $a_{i1}x_1 + \cdots + a_{in}x_n + w_i = b_i$ and $w_i \geq 0$. After this conversion, the LP and its dual are:

$$\begin{array}{ll} \text{maximize } [c0]^T [xw] & \text{minimize } y^T b \\ \text{subject to } [AI][xw] = b,\ [xw] \geq 0 & \text{subject to } y^T [AI] \geq [c0]^T \end{array} \tag{2}$$

where

- $[c0]$ denotes the vector c followed by m zeros,
- $[xw]$ denotes the vector consisting of the n original variables x followed by the m slack variables w, and
- $[AI]$ denotes the $m \times (n+m)$ matrix obtained by concatenating A with an $m \times m$ identity matrix on the right.

Note that $c^T x = [c0]^T [xw]$, so the primal objective function is unchanged, and the dual constraints $y^T A \geq c^T$ and $y \geq 0$ are now both captured by $y^T [AI] \geq [c0]^T$.

We will give the geometric intuition behind this later.

2 Optimization vs Decision

We can also turn an optimization LP into a decision problem by imposing a new constraint on the objective function. The problem then becomes a question of the feasibility of the new set of constraints.

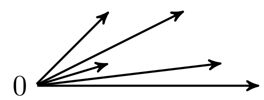
3 Farkas's Lemma

Having equality constraints and feasibility versions of our LPs will allow us to apply *Farkas's lemma*, which is the core of the strong duality result. We will state Farkas's lemma and give the geometric intuition behind it, followed by a proof sketch.

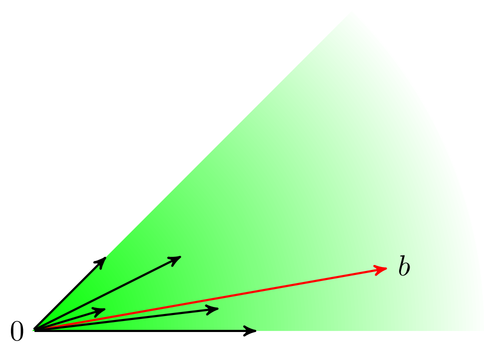
Lemma 3.1 (Farkas's lemma (1902))**.** *Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Exactly one of the following two systems is feasible:*

$$Ax = b,\ x \geq 0 \qquad\qquad y^T A \geq 0,\ y^T b < 0. \tag{3}$$

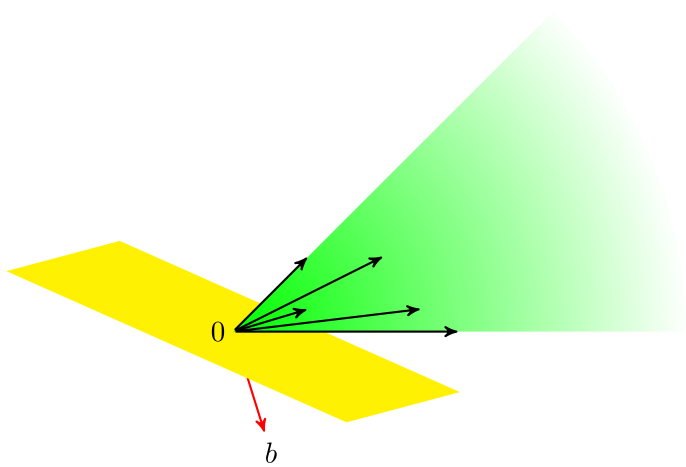
The formal statement of Farkas's lemma may seem a bit mysterious, but here is the geometric intuition behind it. Consider the columns of A as a collection of n vectors in $\mathbb{R}^m$.



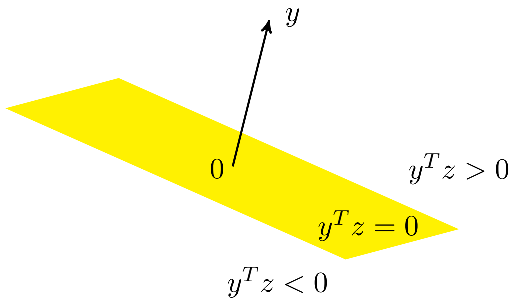
The set of all nonnegative linear combinations of these vectors is the set $\{Ax \mid x \geq 0\}$. Geometrically, this forms a closed convex set in $\mathbb{R}^m$, which we will call the *positive cone* generated by the vectors. To say that the system $Ax = b$, $x \geq 0$ is feasible says that b lies in this positive cone.



If the system is not feasible, then b lies outside the cone. In this case, it can be shown that there exists a hyperplane that separates b from the cone.



A hyperplane in $\mathbb{R}^m$ is subspace of dimension $m-1$ consisting of all vectors orthogonal to some vector y. It partitions $\mathbb{R}^m$ into three disjoint regions distinguished by the inner product with y. The inner product of two nonzero vectors is positive, negative, or zero according as the angle between them is acute (less than $\pi/2$), obtuse (greater than $\pi/2$), or right (exactly $\pi/2$), respectively. The points z for which $y^T z > 0$ lie on the same side of the hyperplane as y and those for which $y^T z < 0$ lie on the opposite side.



A feasible point y of the system $y^T A \geq 0$, $y^T b < 0$ determines such a hyperplane. The condition $y^T A \geq 0$ says that the columns of A form acute or right angles with y, therefore lie on the same side of the hyperplane as y or on the hyperplane itself. It is easily shown that this is also true of all elements in the positive cone they generate. The condition $y^T b < 0$ says that b forms an obtuse angle with y, therefore lies on the opposite side.

To prove Farkas's lemma, we need to show that exactly one of the systems (3) is feasible. Certainly they cannot both be feasible, because then we would have $y^T Ax = y^T b < 0$ and $y^T Ax \geq 0$, a contradiction.

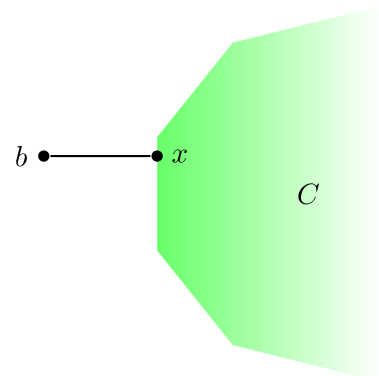
To show that at least one is feasible, suppose the left-hand system of (3) is not feasible. Then b lies outside the positive cone generated by the columns of A. Now we wish to construct a hyperplane separating b from this positive cone.

In fact, it is it possible to construct an affine hyperplane (a hyperplane perhaps translated away from the origin) separating any nonempty closed convex set $C \subseteq \mathbb{R}^m$ from any point $b \notin C$. We use the *Weierstrass extreme value theorem*, which states that a continuous real-valued function f on a compact (closed and bounded) subset of $\mathbb{R}^m$ achieves a minimum value; that is, there exists a point $x \in C$ such that

$$f(x) = \inf_{z \in C} f(z).$$

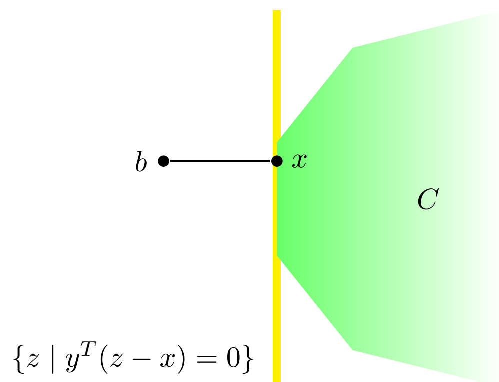
Let $v \in C$ be arbitrary. Applying this theorem to the continuous function $\|b - x\|$ on the closed and bounded set $C \cap \{x \mid \|b - x\| \leq \|b - v\|\}$, we see that there is a point $x \in C$

minimizing the distance to b.



In fact, the point x is unique, although we do not need to know that for our application.

The vector $y = x - b$ is the desired normal vector. In general we would have to translate the hyperplane $\{z \mid y^T z = 0\}$ so the resulting affine hyperplane $\{z \mid y^T(z - x) = 0\}$ would contain x, but it turns out that in our application this is not necessary, as the hyperplane already contains x.



Let us show that the hyperplane separates C from b. Let $x' \in C$ be arbitrary. For any $\varepsilon > 0$, since C is convex, $\varepsilon x' + (1-\varepsilon)x \in C$. Using the fact that $\|c+d\|^2 = \|c\|^2 + \|d\|^2 + 2c^T d$,

$$\begin{aligned}
\|y\|^2 &= \|x - b\|^2 \\
&\le \|(\varepsilon x' + (1-\varepsilon)x) - b\|^2 = \|x - b + \varepsilon(x' - x)\|^2 = \|y + \varepsilon(x' - x)\|^2 \\
&= \|y\|^2 + \|\varepsilon(x' - x)\|^2 + 2y^T\varepsilon(x' - x) \\
&= \|y\|^2 + \varepsilon^2\|x' - x\|^2 + 2y^T\varepsilon(x' - x),
\end{aligned}$$

so

$$y^T(x' - x) \ge -\frac{\varepsilon}{2}\|x' - x\|^2,$$

and taking the limit as $\varepsilon \to 0$ gives $y^T(x' - x) \ge 0$. But since both 0 and $2x$ lie in C,

$$-y^T x = y^T(0 - x) \ge 0 \qquad\qquad y^T x = y^T(2x - x) \ge 0,$$

therefore $y^T x = 0$, so $y^T x' \ge 0$. As for b, we have $y^T b = y^T(b - x) = -y^T y < 0$.

4 Strong Duality

Now we can give a geometric interpretation to the construction of (2) from (1) in §1. To say that there exists an x such that $Ax \leq b$ and $x \geq 0$ says that there is a point c in the positive cone $\{Ax \mid x \geq 0\}$ such that $c \leq b$; in other words, there is a point c in the positive cone and a further nonnegative vector z such that $c + z = b$. This is the same as adding the unit standard basis vectors to the columns of A to get $[AI]$ and asking whether b is in the positive cone generated by the columns of $[AI]$. That is the system (2).

To show the strong duality result, we start with the primal and dual systems (1). We know from weak duality that $\max c^T x \leq \min y^T b$. To show that they are equal, we show that for any d, if $\max c^T x < d$, then $\min y^T b < d$ as well. That is, if the maximization problem with the extra constraint $c^T x \geq d$ is infeasible, then $\min y^T b < d$.

Let us assume that the system $Ax \leq b$, $x \geq 0$ is feasible and add the extra constraint $c^T x \geq d$, or equivalently, $-c^T x \leq -d$. This can be represented by taking A' to be A with the extra row $-c^T$, b' to be b with the extra element $-d$, and taking the constraints to be $A'x \leq b'$.

Now we add slack variables w as in §1 to get a system with equality constraints

$$[A'I][xw] = b' \qquad\qquad [xw] \geq 0.$$

If this system is infeasible, Farkas's lemma says that there exist a vector of m values y and a scalar z such that

$$[yz]^T[A'I] \geq 0 \qquad\qquad [yz]^T b' < 0,$$

where $[yz]$ is the vector y with z appended. Equivalently,

$$y^T A - zc^T \geq 0 \qquad\qquad y, z \geq 0 \qquad\qquad y^T b - zd < 0.$$

It cannot be that $z = 0$, because otherwise $y^T A \geq 0$, $y \geq 0$, $y^T b < 0$ would be feasible, which would mean that the system $Ax \leq b$, $x \geq 0$ would be infeasible by Farkas's lemma. But this contradicts our assumption. Thus $z > 0$. We have

$$y^T A \geq zc^T \qquad\qquad y, z \geq 0 \qquad\qquad y^T b < zd.$$

Let y/z be the vector y scaled by the value z. Then

$$(y/z)^T A \geq c^T \qquad\qquad y/z \geq 0 \qquad\qquad (y/z)^T b < d,$$

so we have a witness that the minimum value of the right-hand system of (1) is less than d.