6.854J / 18.415J Advanced Algorithms
Fall 2008

18.415/6.854 Advanced Algorithms	September 19, 2001

Lecture 3

Lecturer: Michel X. Goemans *Scribe: Abhinav Kumar, Nodari Sitchinava*

1 Proof of Farkas' Lemma

Theorem 1 **[Farkas' Lemma]** *Either*

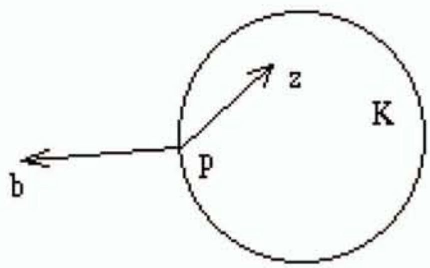
1. $Ax = b$, $x \geq 0$ *has a solution, or*
2. $A^T y \geq 0$ *and* $y^T b < 0$ *has a solution,*

but not both.

The reason that 1 and 2 cannot both occur is that $(y^T A)x = y^T b$, so if $y^T A$ is non-negative and x is non-negative, then $y^T b$ can't be negative.

To prove Farkas' Lemma we need the Projection Theorem:

Theorem 2 *Let K be a closed, convex and non-empty set in $\mathbb{R}^n$, and $b \in \mathbb{R}^n$, $b \notin K$. Define projection p of b onto K to be $x \in K$ such that $||b - x||$ is minimized. Then for all $z \in K$: $(b-p)^T(z-p) \leq 0$.*



Proof of Farkas' Lemma: Assume $Ax = b, x \geq 0$ is not feasible. Let $K = \{Ax : x \geq 0\}$. Therefore, $b \notin K$. Let $p = Aw, w \geq 0$ be the projection of b onto K. Then we know that

$$(b - Aw)^T(Ax - Aw) \leq 0 \text{ for all } x \geq 0 \tag{1}$$

Define $y = p - b = Aw - b$. Therefore,

$$(x - w)^T A^T y \geq 0 \text{ for all } x \geq 0 \tag{2}$$

Let e_i be the $n \times 1$ vector that has 1 in its i'th component and 0 everywhere else. Take $x = w + e_i$. Therefore, $x - w = e_i$, and by (2),

$$e_i^T A^T y \geq 0 \Rightarrow (A^T y)_i \geq 0 \text{ for all } i$$

Thus since each element of $A^T y$ is non-negative, $A^T y \geq 0$.

Now, $y^Tb = y^T(p-y) = y^Tp - y^Ty$. ¿From (1) if $x = 0$,

$$\begin{aligned}(b-Aw)^T(Ax-Aw) &= (b-Aw)^T(-Aw)\\ &= -y^T(-Aw)\\ &= y^TAw\\ &= y^Tp \leq 0\end{aligned}$$

and

$$y^Tp - y^Ty \leq -y^Ty < 0$$

The last inequality comes from the fact that $y = b - p, b \notin K$, so $b - p \neq 0 \Rightarrow y^Ty > 0$ □

Theorem 3 **[Another variant of Farkas' Lemma]** *Either*

1. $Ax \leq b$ *has a solution, or*
2. $A^Ty = 0, b^Ty < 0, y \geq 0$ *has a solution,*
   but not both (for then we would have $0 = y^TAx \leq y^Tb < 0$.*)*

2 Duality

Consider an LP P in the standard form (we call this LP the primal). We can write a "dual" LP D as follows:

Primal P: $z^* = \min c^Tx$
subj to
$Ax = b$
$x \geq 0$

Dual D: $w^* = \max b^Ty$
subj to
$A^Ty \leq c$

Weak duality states the following.

Theorem 4 **[Weak Duality]** *Let x be feasible in P, and let y be feasible in D. Then*

$$c^Tx \geq b^Ty$$

Proof of Theorem 4:

$$\begin{aligned}c^Tx - b^Ty &= x^Tc - x^TA^Ty\\ &= x^T(c - A^Ty)\\ &\geq 0,\end{aligned}$$

since x and $c - A^Ty$ both have nonnegative coordinates.

□

The following three cases are possible for an LP:

Primal	Dual
1) infeasible ($z^* = +\infty$)	1') infeasible ($w^* = -\infty$)
2) unbounded ($z^* = -\infty$)	2') unbounded ($w^* = +\infty$)
3) finite (z^* = finite real number)	3') finite (w^* = finite real number)

Then 2 $\Rightarrow$ 1' because if the dual were feasible, any value $b^T y$ for the dual would be a lower bound for the primal, which could therefore not be unbounded. Similarly 2' $\Rightarrow$ 1. Note that we can have 1 and 1' occurring simultaneously.

Theorem 5 [Strong duality] *If P or D is feasible then $z^* = w^*$.*

Proof of Theorem 2: It suffices to treat the case when the primal is feasible, because the primal and dual are interchangeable. So assume P is feasible. If P is unbounded then weak duality implies that D is infeasible, and then $z^* = w^* = -\infty$. So from now on assume that the primal is finite.

Claim 6 *There exists a solution of dual of value at least z^*, i.e.,*

$$\exists y : A^T y \leq c, b^T y \geq z^*$$

Proof of Claim 3: We wish to prove that there is a y satisfying

$$\begin{pmatrix} A^T \\ -b^T \end{pmatrix} y \leq \begin{pmatrix} c \\ -z^* \end{pmatrix}.$$

Assume the claim is wrong. Then the variant of Farkas' Lemma implies that the LP

$$\begin{aligned} \begin{pmatrix} A & -b \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} &= 0 \\ \begin{pmatrix} c^T & -z^* \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} &< 0 \\ x \in \mathbb{R}^n, \lambda \in \mathbb{R}, \quad x, \lambda &\geq 0 \end{aligned}$$

has a solution. That is, there exist nonnegative x, λ with

$$\begin{aligned} Ax - b\lambda &= 0 \\ c^T x - z^* \lambda &< 0 \end{aligned}$$

Case 1: $\lambda > 0$. Then $A(\frac{x}{\lambda}) = b, \quad c^T(\frac{x}{\lambda}) < z^*$. This contradicts the minimality of z^* for the primal, hence this case cannot occur.

Case 2: $\lambda = 0$. Then $Ax = 0, \quad c^T x < 0$. Take any feasible solution $\hat{x}$ for P. Then for every $\mu \geq 0$, $\hat{x} + \mu x$ is feasible for P, since

a) $\hat{x} + \mu x \geq 0$ because $\hat{x} \geq 0, x \geq 0, \mu \geq 0$.

b) $A(\hat{x} + \mu x) = A\hat{x} + \mu A x = b + \mu \cdot 0 = b$.

But $c^T(\hat{x} + \mu x) = c^T \hat{x} + \mu c^T x \to -\infty$ as $\mu \to \infty$. This contradicts the assumption that the primal has finite solution.

□

The above claim shows that if P or D is finite then the other is too, and the optimums are equal ($z^* \geq w^*$ is weak duality and the claim shows $w^* \geq z^*$.) This concludes the proof of the strong duality theorem. □

3 Complementary Slackness

Consider the following primal LP.

$$
\begin{aligned}
\min\ & c^T x \\
Ax &= b \\
x &\geq 0
\end{aligned}
$$

We write the dual as follows:

$$
\begin{aligned}
\max\ & b^T y \\
A^T y + s &= c \\
s &\geq 0, \quad y \in \mathbb{R}^m, s \in \mathbb{R}^n
\end{aligned}
$$

Theorem 7 *Let x be feasible for the primal, and y be feasible for the dual. Then x is optimal for P and y is optimal for D if and only if $x_j s_j = 0$ for all j.*

Proof: We have

$$
\begin{aligned}
c^T x - b^T y &= x^T c - x^T A^T y \\
&= x^T (c - A^T y) \\
&= x^T s
\end{aligned}
$$

When both x and y are optimal, the above difference must be zero, and conversely, if the difference is zero, both must be optimal by weak duality. But since x, s are nonnegative, $x^T s$ is zero if and only if $x_j s_j = 0$ for all j. $\square$.

So, to prove that a solution to an LP is optimal, all we need to do is to give an x and a (y, s) and show that both are feasible and the complementary slackness condition is satisfied.

4 Size of a linear program

Let's think about how we encode the LP. We can use binary encoding to give the entries of A, b, c, that defines the LP in standard form. For an integer k, it takes $size(k) = 1 + \lceil \log_2(|k| + 1) \rceil$ bits to encode k. So,

$$
size(LP) = \sum_{i,j} size(a_{ij}) + \sum_j size(c_j) + \sum_i size(b_i)
$$

A polynomial-time algorithm for linear programming is an algorithm whose worst-case running time is bounded by a polynomial in the size of the input LP.