NOTES ON FARKAS' LEMMA AND THE STRONG DUALITY THEOREM FOR LINEAR PROGRAMMING

Let A be an $m \times n$ real matrix. We work throughout with column vectors $x \in \mathbf{R}^n$ and $y \in \mathbf{R}^m$.

Farkas' Lemma. Let $b \in \mathbf{R}^m$. Either there exists $x \in \mathbf{R}^n$ such that $x \geq 0$ and $Ax = b$, or there exists $y \in \mathbf{R}^m$ such that $y^{\mathrm{t}} A \geq 0$ and $y^{\mathrm{t}} b = -1$.

In geometric language, Farkas' Lemma says that *either* b is in the positive cone made by the basis vectors $Ae_1, \ldots, Ae_n$, where $e_1, \ldots, e_n$ is the standard basis of $\mathbf{R}^n$, *or* there is a hyperplane $H = \{v \in \mathbf{R}^n : y^t v = 0\}$ such that every point of the cone is on the positive side of H (algebraically, $y^t A x \geq 0$), and b is on the negative side of H (algebraically, $y^t b = -1$). It is obvious that at most one of these possibilities can hold.

Proof of Farkas' Lemma. Let $C = \{Ax : x \in \mathbf{R}^n, x \geq 0\}$ and suppose that the 'either' case fails to hold, so $b \notin C$. Let B be the closed ball of radius R about b, where $R \geq ||b||$, so $B \cap C \neq \varnothing$. Since C is closed and B is compact, there is a closest vector $w \in B \cap C$ to b. If $u \notin B$ then $||b - u|| \geq R$, whereas $||b - w|| \leq R$, so w is a closest vector in C to b.

Take any $v \in C$ and consider the line segment joining v to w. Since C is convex, this segment lies in C. Therefore for any α such that $0 \leq \alpha \leq 1$, we have

$$\begin{aligned} ||w - b||^2 &\leq ||\alpha v + (1-\alpha) w - b||^2 \\ &= ||\alpha(v - w) + (w - b)||^2 \\ &= \alpha^2 ||v - w||^2 + 2\alpha \langle v - w, w - b\rangle + ||w - b||^2. \end{aligned}$$

Hence, taking α to be small and positive, we see that $\langle v - w, w - b\rangle \geq 0$.

Set $z = w - b$. Setting $v = 0$ we get $\langle w, z\rangle \leq 0$, and so

$$\langle b, z\rangle = \langle b - w, z\rangle + \langle w, z\rangle = -||w - b||^2 + \langle w, z\rangle < 0.$$

Hence there exists $\gamma < 0$ such that

$$\langle v, z\rangle \geq \langle w, z\rangle > \gamma > \langle b, z\rangle$$

for all $v \in C$. Fix $v \in C$. Since $\langle \lambda v, z\rangle > \gamma$ for all $\lambda \geq 0$ we have $\langle v, z\rangle > \gamma/\lambda$ for all $\lambda > 0$. Taking λ large we see that $\langle v, z\rangle \geq 0$. Therefore $y = z/|\langle z, b\rangle|$ satisfies the required conditions for the 'or' case. $\square$

Date: February 8, 2024.

Variant Farkas' Lemma. For the application to the strong duality theorem we need a slightly different version of Farkas' Lemma.

Lemma 1. *Let $b \in \mathbf{R}^m$. Either there exists $x \in \mathbf{R}^n$ such that $Ax \leq b$, or there exists $y \in \mathbf{R}^m$ such that $y \geq 0$, $y^{\mathrm{t}} A = 0$ and $y^{\mathrm{t}} b = -1$.*

This lemma also has a geometric interpretation, although it maybe takes a bit more effort to see. Note that if $y \geq 0$ (and $y \neq 0$) then the hyperplane $H = \{v \in \mathbf{R}^n : y^t v = 0\}$ contains no $v \in \mathbf{R}^n$ such that $v \geq 0$ or $v \leq 0$, except for $v = 0$. Say, just for this paragraph, that such hyperplanes are *tilted*. Note that a non-tilted hyperplane always contain some $\alpha e_i + \beta e_j$ with $i < j$ and $\alpha, \beta > 0$. So if H is a hyperplane such that $v \not\leq b$ for all $v \in H$, then v is tilted. Lemma 1 says that *either* there exists $v \in \operatorname{im} A$ such that $v \leq b$, *or* $\operatorname{im} A$ is contained in a tilted hyperplane H that has b on its negative side, and so $Ax \not\leq b$ for all $x \in \mathbf{R}^n$. Again it is obvious that at most one of these possibilities can hold.

Algebraic proof of equivalence of Farkas' Lemma and Lemma 1. Suppose that Farkas' Lemma holds. If the 'or' case of Lemma 1 fails to hold then there is no $y \in \mathbf{R}^m$ such that

$$y^{\mathrm{t}} \begin{pmatrix} A & I_m \end{pmatrix} \geq 0$$

and $y^{\mathrm{t}} b = -1$. Hence, by Farkas' Lemma, there exists $x \in \mathbf{R}^n$ and $z \in \mathbf{R}^m$ such that that $x \geq 0$, $z \geq 0$ and

$$\begin{pmatrix} A & I_m \end{pmatrix} \begin{pmatrix} x \\ z \end{pmatrix} = b$$

Therefore $Ax \leq b$ and the 'either' case of Lemma 1 holds.

Suppose that Lemma 1 holds. If the 'either' case of Farkas' Lemma fails to hold then there is no $x \in \mathbf{R}^n$ such that $x \geq 0$ and

$$\begin{pmatrix} A & -b \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = 0.$$

Let $c = (0, \ldots, 0, 1) \in \mathbf{R}^{n+1}$. By scaling, and transposing, it follows that there is no $w \in \mathbf{R}^{n+1}$ such that $w \geq 0$,

$$w^t \begin{pmatrix} A^t \\ -b^t \end{pmatrix} = 0$$

and $-w^t c = -1$. Hence, by Lemma 1, there exists $y \in \mathbf{R}^m$ such that $y \geq 0$ and

$$\begin{pmatrix} A^t \\ -b^t \end{pmatrix} y \leq c.$$

Thus $A^t y \leq 0$ and $b^t y \geq 1$. By scaling we may assume that $b^t y = 1$, as required for the 'or' case of Farkas' Lemma.

Primal and dual problems. In the remainder of this note we consider the following two standard optimization problems. Let $b \in \mathbf{R}^m$ and let $c \in \mathbf{R}^n$.

Primal. Maximize $c^t x$ for $x \in \mathbf{R}^n$ subject to $x \geq 0$, $Ax \leq b$.

Dual. Minimize $y^t b$ for $y \in \mathbf{R}^m$ subject to $y \geq 0$, $y^t A \geq c^t$.

As usual, we say that the primal problem is *feasible* if there exists $x \in \mathbf{R}^n$ such that $x \geq 0$ and $Ax \leq b$. We say that the primal problem is *unbounded* if there exist feasible $x \in \mathbf{R}^n$ such that $c^t x$ takes arbitrarily large values. We define *feasible* and *unbounded* for the dual problem similarly.

Theorem 2 (Weak Duality). *If x is feasible for the primal problem and y is feasible for the dual problem then $c^t x \leq y^t b$.*

Proof. From $Ax \leq b$ and $c^t \leq y^t A$ we get $c^t x \leq y^t A x \leq y^t b$, as required. □

In particular if the primal problem is feasible then the dual problem is bounded, and if the dual problem is feasible then the primal problem is bounded. The converses to these statements do not hold, because it is possible that both primal and dual problems are infeasible (and so, immediately from the definition, they are both bounded): for example, take

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad c = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Then for any $x \in \mathbf{R}^2$ such that $x \geq 0$ we have $Ax \not\leq b$, and for any $y \in \mathbf{R}^2$ such that $y \geq 0$ we have $y^t A \not\geq c^t$. The theorem below gives the strongest possible converse.

Theorem 3. *The primal problem is feasible and bounded if and only if the dual problem is feasible and bounded.*

Proof. By the Weak Duality Theorem, if the primal problem is unbounded then the dual problem is infeasible. Suppose that the primal problem is infeasible. So there exists no $x \in \mathbf{R}^n$ such that $x \geq 0$ and $Ax \leq b$. Hence there exists no $x \in \mathbf{R}^n$ such that

$$\begin{pmatrix} A \\ -I_n \end{pmatrix} x \leq \begin{pmatrix} b \\ 0 \end{pmatrix}.$$

By Lemma 1 there exist $y \in \mathbf{R}^m$ and $z \in \mathbf{R}^n$ such that $(y^t \; z^t) \geq 0$,

$$(y^t \;\; z^t) \begin{pmatrix} A \\ -I_n \end{pmatrix} = 0$$

and $y^t b = -1$. If the dual problem is infeasible we are done, so we may suppose that there exists a feasible $y_\star \in \mathbf{R}^m$ for the dual problem. We have $(y_\star^t + \lambda y^t)A \geq c^t + \lambda z^t \geq c^t$ and $(y_\star^t + \lambda y^t)b = y_\star^t b - \lambda$. Taking λ large we see that the dual problem is unbounded. The converse has a very similar proof. □

This motivates restricting to the case where both the primal problem and the dual problem are feasible.

Theorem 4 (Strong Duality Theorem)**.** *If both the primal and dual problems are feasible then they have the same optimal value.*

We prove this theorem by extending the argument used to prove Theorem 3.

Proof of Strong Duality Theorem. Let $\tau_P \in \mathbf{R}$ be the optimal value of the primal problem and let $\tau = \tau_P + \varepsilon$. Since there exists no $x \in \mathbf{R}^n$ such that $Ax \leq b$, $c^t x \geq \tau$, there exists no $x \in \mathbf{R}^n$ such that

$$\begin{pmatrix} A \\ -I_n \\ -c^t \end{pmatrix} x \leq \begin{pmatrix} b \\ 0 \\ -\tau \end{pmatrix}.$$

By Lemma 1 there exist $y \in \mathbf{R}^m$, $z \in \mathbf{R}^n$ and $\alpha \in \mathbf{R}$ such that $(y^t\ z^t\ \alpha) \geq 0$,

$$(y^t \quad z^t \quad \alpha) \begin{pmatrix} A \\ -I_n \\ -c^t \end{pmatrix} = 0$$

and $y^t b - \alpha\tau < 0$. (It is more convenient to have an inequality here to permit scaling later.) Thus $y^t A = z^t + \alpha c^t$ and $y^t b < \alpha\tau$.

Suppose that $\alpha = 0$. Then, since $y^t A = z^t \geq 0$ and $y^t b = -1$, it follows as in the proof of Theorem 3 that the dual problem is either infeasible or unbounded. This contradicts the Weak Duality Theorem since, by hypothesis, both problems are feasible. Therefore $\alpha \neq 0$ and by scaling we may assume that $\alpha = 1$. So $y^t A \geq c^t$ and $y^t b < \tau$. Hence if $\tau_D \in \mathbf{R}$ is the optimal value of the dual problem then $\tau_D < \tau = \tau_P + \varepsilon$. By the Weak Duality Theorem it follows that $\tau_P \leq \tau_D < \tau_P + \varepsilon$. Since ε was arbitrary we have $\tau_P = \tau_D$, as required. $\square$

In practice it suffices to find $x \in \mathbf{R}^n$ and $y \in \mathbf{R}^m$ such that $c^t x \geq y^t b$. Then by the Weak Duality Theorem $c^t x = y^t b$ is the common optimal value of the primal and dual problems.