LINEAR PROGRAMMING DUALITY AND THE MINIMAX THEOREM

DAVID RAHMAN

1. The Primal and Dual

A *linear program* is an optimization problem with a linear objective and linear constraints. Here I will only consider problems with finitely many choice variables and constraints. As we shall see, every linear program has another intimately related one—they come in pairs.

Fix (A, b, c) such that $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$. The *primal* is the linear program

$$V = \sup_{x\geq \mathbf{0}}\{cx : Ax \leq b\},$$

and the *dual* is the linear program

$$W = \inf_{y\geq \mathbf{0}}\{yb : yA \geq c\}.$$

A linear program is in "standard form" if it looks like a primal or a dual for some (A, b, c). It is easy to see that every linear program is expressible in standard form. A linear program is called *infeasible* if its constraint set is empty. By convention, $V = -\infty$ if the primal is infeasible and $W = +\infty$ if the dual is infeasible. A linear program is *feasible* if its constraint set is nonempty. In this case, any feasible vector is called a *feasible solution.*[1]

Theorem 1 (Weak Duality)**.** *$V \leq W$ for any triple (A, b, c).*

Proof. If the primal is infeasible then $V = -\infty$, so necessarily $V \leq W$ regardless of W, and if the dual is infeasible then $W = +\infty$; again $V \leq W$. Now assume that neither the primal nor the dual is infeasible. That is, there exist vectors $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ such that $x, y \geq \mathbf{0}$, $Ax \leq b$ and $yA \geq c$. This implies that $(b-Ax) \geq \mathbf{0}$. Multiplying by $y \geq \mathbf{0}$ gives $y(b-Ax) \geq 0$, or $yb \geq yAx$. Similarly, multiplying $(yA - c) \geq \mathbf{0}$ by $x \geq \mathbf{0}$ gives $(yA - c)x \geq 0$, or $yAx \geq cx$. Collecting these inequalities gives

$$cx \leq yAx \leq yb$$

for every feasible pair (x, y). Therefore, the primal and dual are both *bounded*:

$$-\infty < cx \leq yb < +\infty \qquad \Rightarrow \qquad V \leq yb < +\infty \quad \text{and} \quad -\infty < cx \leq W$$

by taking the supremum of the LHS of $cx \leq yb$ with respect to feasible x given yb and the infimum of the RHS of $cx \leq yb$ with respect to feasible y given x. Taking the infimum of the RHS of $V \leq yb$ with respect to feasible y or the supremum of the LHS of $cx \leq W$ with respect to feasible x finally gives $V \leq W$, and the theorem is proved. □

Date: January 29, 2025.

Department of Economics, University of Minnesota.

[1]A feasible solution does not necessarily solve the linear program. For instance, $x = 1$ is a feasible solution of the linear program $\max_x\{x : x \leq 2\}$, but it is obviously solved by $x = 2$.

The proof of Theorem 1 above suggests a classification of linear programs into three kinds: infeasible (I), (feasible and) bounded (B), (feasible and) unbounded (U). We will show that the only possibilities for a primal and its dual are those described in Table 1 below.

		Dual I	*Dual* B	*Dual* U
Primal	I	✓		✓
	B		✓	
	U	✓		

TABLE 1. Joint classification of a primal and its dual

Proposition 2. *If the primal is infeasible then the dual is either infeasible or unbounded.*

Proof. If the primal is infeasible then there does not exist a vector $x \geq \mathbf{0}$ such that $Ax \leq b$. By Farkas' Lemma, there exists a vector $y \geq \mathbf{0}$ such that $yA \geq 0$ and $yb < 0$. If the dual is infeasible, the proposition is proved, so suppose that the dual is feasible. For any feasible dual solution z and any $\lambda \geq 0$, the vector $z + \lambda y$ is also feasible, since $z + \lambda y \geq \mathbf{0}$ and $(z + \lambda y)A = zA + \lambda yA \geq c + \lambda yA \geq c$. Moreover, $(z + \lambda y)b = zb + \lambda yb$ is unbounded: since $yb < 0$, for any bound $K \in \mathbb{R}$ there exists λ large enough that $zb + \lambda yb < K$. □

The contrapositive of Proposition 2 is this: if the dual is bounded then its primal is feasible. Therefore, both primal and dual are feasible, so by the proof of Theorem 1 they are bounded, too. Since the dual of the dual is the primal, as long as either the primal or the dual is bounded, both are. Moreover, their values coincide, as the next result shows.

Theorem 3 (Strong Duality). *Unless both the primal and the dual are infeasible, $V = W$. Moreover, both the primal and the dual have optimal solutions whenever either is bounded.*

Proof. If the primal is feasible but unbounded then $V = +\infty$. If its dual were feasible then there would exist $y \geq \mathbf{0}$ such that $yA \geq c$. But this would contradict Theorem 1, since it would imply that $W \leq yb < +\infty = V$. Therefore, the dual is infeasible, and $V = W$.

If the primal is bounded then, by the argument immediately following the proof of Proposition 2, the dual is bounded, too. Therefore, both primal and dual are feasible. By Farkas' Lemma, primal feasibility implies that there is no vector $u \geq \mathbf{0}$ such that $uA \geq \mathbf{0}$ and $ub < 0$, whereas dual feasibility implies that there is no vector $v \geq \mathbf{0}$ such that $Av \leq \mathbf{0}$ and $cv > 0$. Call these the *primal and dual feasibility conditions*.

Given that both primal and dual are feasible, we must show that their values coincide, that is, there exist vectors $x, y \geq \mathbf{0}$ such that $Ax \leq b$, $yA \geq c$ and $yb \leq cx$. (By Theorem 1, the inequality $cx \leq yb$ already holds.) Let us now collect these inequalities into a single system. Define $z = (x, y') \in \mathbb{R}^{n+m}$, $d = (b, -c', 0) \in \mathbb{R}^{m+n+1}$ and $B \in \mathbb{R}^{(m+n+1)\times(n+m)}$ by

$$B = \begin{bmatrix} A & 0 \\ 0 & -A' \\ -c & b' \end{bmatrix}.$$

By construction, there exists a vector $z \geq 0$ such that $Bz \leq d$ if and only if the primal and dual values coincide at x and y, respectively. By Farkas' Lemma, either this system of inequalities holds or there is a vector $w \geq \mathbf{0}$ such that $wB \geq \mathbf{0}$ and $wd < 0$. Writing $w = (u, v', \lambda) \in \mathbb{R}^{m+n+1}$, where $\lambda \in \mathbb{R}$, the inequalities $wB \geq \mathbf{0}$ and $wd < 0$ disaggregate into $uA - \lambda c \geq \mathbf{0}$, $-v'A' + \lambda b' \geq \mathbf{0}$, and $ub - v'c' < 0$, or equivalently

$$uA \geq \lambda c, \qquad Av \leq \lambda b, \qquad ub < cv.$$

We will now show that there does not exist a vector $w \geq \mathbf{0}$ that satisfies these inequalities. To see why, suppose that $w = (u, v', \lambda)$ satisfies the system. If $\lambda > 0$ then, dividing by λ and substituting $\hat{u} = u/\lambda \geq \mathbf{0}$, $\hat{v} = v/\lambda \geq \mathbf{0}$ into the system of inequalities above gives

$$\hat{u}A \geq c, \qquad A\hat{v} \geq b, \qquad \hat{u}b < c\hat{v}.$$

This contradicts Theorem 1, as it states that there is a primal feasible solution $\hat{u}$ and a dual feasible solution $\hat{v}$ such that weak duality fails. Therefore, $\lambda > 0$ leads to a contradiction. If instead $\lambda = 0$, the inequalities above become

$$uA \geq \mathbf{0}, \qquad Av \geq \mathbf{0}, \qquad ub < cv.$$

But this contradicts the primal and dual feasibility conditions of the previous page, since they stipulate that $ub \geq 0 \geq cv$. Therefore, the alternative to these inequalities must hold, in other words, there exist primal and dual feasible solutions whose values coincide. By weak duality, these variables are optimal for the respective linear programs in which they are feasible. □

Theorem 3 and its proof facilitate the analysis of linear programs substantially. To find an optimal solution, it is enough to find feasible primal and dual solutions with the same value. Moreover, to guarantee strong duality, that is, $V = W$, it is enough to establish that either the primal or the dual is feasible. Example 4 below illustrates the only case in which the value of the primal can differ from the value of the dual, namely when they are both infeasible.

Example 4. Consider the following linear programs:

$$V = \sup_{x \in \mathbb{R}^2} \{x_1 - x_2 : x_1 + x_2 = 1,\ x_1 + x_2 = -1\}, \quad W = \inf_{y \in \mathbb{R}^2} \{y_1 - y_2 : y_1 + y_2 = 1,\ y_1 + y_2 = -1\}.$$

It is easy to see that the dual of the linear program on the left above is the linear program on the right.[2] Clearly, the primal and dual feasible sets coincide. However, both sets are obviously empty, therefore $-\infty = V < W = +\infty$.

Theorem 5 (Complementary Slackness)**.** *The feasible* (x^*, y^*) *satisfies*

$$(y^*A - c)x^* = 0 \qquad \textit{and} \qquad y^*(Ax^* - b) = 0$$

if and only if (x^*, y^*) *is optimal for the primal and the dual, respectively.*

[2]Yes, these linear programs are not in standard form. The purported primal can be turned into an equivalent linear program in standard form, and its dual is easily seen to be equivalent to the purported dual above. Alternatively, at the end of Section 2 it will be seen how to take the dual of a nonstandard linear program.

Proof. If the condition above holds then $cx^* = y^*Ax^* = y^*b$, that is, x^*, y^* are feasible primal and dual solutions with the same value, so by Theorem 3 they are optimal solutions. Conversely, given two optimal solutions, again by Theorem 3, they must have the same value, so $cx^* = y^*b$. Therefore, $(y^*A - c)x^* = y^*(Ax^* - b)$. By feasibility, $Ax^* - b \leq 0$ and $y^*A - c \geq 0$, so, since $x^*, y^* \geq \mathbf{0}$, it follows that $(y^*A - c)x^* \geq 0 \geq y^*(Ax^* - b)$, which implies $y^*Ax^* - cx^* = 0 = y^*Ax^* - y^*b$. □

The equations of Theorem 5, called *complementary slackness*, are often described as follows. By feasibility, $x^* \geq \mathbf{0}$ and $y^*A - c \geq \mathbf{0}$, so $(y^*A - c)x^* = \sum_j ((y^*A)_j - c_j)x_j^* = 0$ requires

$$((y^*A)_j - c_j)x_j^* = 0 \qquad \text{for } j = 1, \ldots, n.$$

In other words, if $x_j^* > 0$ then $(y^*A)_j = c_j$, and if $(y^*A)_j > c_j$ then $x_j^* = 0$. A corresponding conclusion applies to the dual equation, of course.

2. Constrained Optimization as a Game

The primal $V = \max\{cx : Ax \leq b, x \geq \mathbf{0}\}$ involves maximizing a linear objective subject to linear constraints. For simplicity, assume that the constraint set is nonempty and the value V is bounded. The Lagrangian approach to this optimization problem is to construct a vector of Lagrange multipliers, call them $y \in \mathbb{R}^m$, so there is one multiplier per constraint, and play the following game of opposing interests. First, player 1, say, chooses a vector $x \in \mathbb{R}^n$ such that $x \geq \mathbf{0}$. Next, after observing player 1's move, player 2 chooses a vector $y \in \mathbb{R}^m$ such that $y \geq \mathbf{0}$. Player 2 then pays player 1 the amount stipulated in the Lagrangian below:

$$L(x, y) = cx + y(b - Ax).$$

If player 1 chooses x and player 2 chooses y then player 2 pays player 1 $L(x, y)$. If player 1 chose $x \geq \mathbf{0}$ that violates the constraint $Ax \leq b$, say because the ith constraint satisfies $(Ax)_i > b_i$, then player 2 can choose $y_i > 0$ increasingly large until $y_i[b_i - (Ax)_i] < K$ for any lower bound $K \in \mathbb{R}$. This way, player 2 can earn $+\infty$ from player 1. Therefore, player 1 will never choose x that violates the constraint $Ax \leq b$. In other words,

$$V = \max_{x \geq \mathbf{0}} \min_{y \geq \mathbf{0}} L(x, y).$$

Similarly, starting with the dual

$$W = \min_{y \geq \mathbf{0}} \{yb : yA \geq c\}$$

and denoting by $x \in \mathbb{R}^n$ the Lagrange multipliers of the dual constraints, it follows that

$$W = \min_{y \geq \mathbf{0}} \max_{x \geq \mathbf{0}} L(x, y).$$

Corollary 6 (Minimax Theorem)**.** *It doesn't matter who goes first in the game above:*

$$\max_{x \geq \mathbf{0}} \min_{y \geq \mathbf{0}} L(x, y) = \min_{y \geq \mathbf{0}} \max_{x \geq \mathbf{0}} L(x, y).$$

In general, a given linear program with some equality constraints would have multipliers that need not be nonnegative. To see this, if there was a constraint of the form $(Ax)_i = b_i$ then its multiplier y_i should have to be allowed to be positive or negative in order for the game to impose the constraint on the player choosing x. More generally, let $x = (x^1, x^2)$ be choice variables such that $x^1 = (x^1_1, \ldots, x^1_{n_1})$ and $x^2 = (x^2_1, \ldots, x^2_{n_2})$, and $c = (c^1, c^2)$ be an objective with $c^1 = (c^1_1, \ldots, c^1_{n_1})$ and $c^2 = (c^2_1, \ldots, c^2_{n_2})$. Let $b = (b^1, b^2)$ such that $b^1 = (b^1_1, \ldots, b^1_{m_1})$ and $b^2 = (b^2_1, \ldots, b^2_{m_2})$ be right-hand side constraints, and $A^{11} \in \mathbb{R}^{m_1 \times n_1}$, $A^{12} \in \mathbb{R}^{m_1 \times n_2}$, $A^{21} \in \mathbb{R}^{m_2 \times n_1}$, and $A^{22} \in \mathbb{R}^{m_2 \times n_2}$ be constraint matrices. Consider the linear program

$$V = \sup_{x \in \mathbb{R}^n} \{cx : [A^{11} \; A^{12}]x = b^1, \; [A^{21} \; A^{22}]x \leq b^2, \; x^2 \geq \mathbf{0}\},$$

where $n = n_1 + n_2$ and $[A^{11} \; A^{12}]$ is the horizontal concatenation of A^{11} and A^{12} (similarly for $[A^{21} \; A^{22}]$). This program has unrestricted choice variables and non-negative ones, equality constraints and inequality constraints. Call it the nonstandard primal. Its Lagrangian is

$$L(x, y) = cx + y^1(b^1 - [A^{11} \; A^{12}]x) + y^2(b^2 - [A^{21} \; A^{22}]x),$$

where the vector of multipliers $y = (y^1, y^2)$ satisfies $y^1 = (y^1_1, \ldots, y^1_{m_1})$ and $y^2 = (y^2_1, \ldots, y^2_{m_2})$ with y^1 unrestricted and $y^2 \geq \mathbf{0}$. Rearranging and multiplying by -1 gives

$$\begin{aligned} M(x, y) = -L(x, y) &= -y^1 b^1 - y^2 b^2 + y^1 A^{11} x^1 + y^1 A^{12} x^2 + y^2 A^{21} x^1 + y^2 A^{22} x^2 - cx \\ &= -yb + (c^1 - y(A^{11}, A^{21}))x^1 + (c^2 - y(A^{12}, A^{22}))x^2, \end{aligned}$$

where $(A^{11}, A^{21}) \in \mathbb{R}^{m \times n_1}$ concatenates A^{11} and A^{21} vertically ($m = m_1 + m_2$); similarly, $(A^{12}, A^{22}) \in \mathbb{R}^{m \times n_2}$. This Lagrangian corresponds to the minimization

$$W = \inf_{y \in \mathbb{R}^m} \{yb : y(A^{11}, A^{21}) = c^1, \; y(A^{12}, A^{22}) \geq c^2, \; y^2 \geq \mathbf{0}\}.$$

Call this the nonstandard dual of the nonstandard primal. Obviously, the Minimax Theorem above applies here, as do the previous duality results. This derivation easily yields the duality of Example 4. Henceforth, we drop standard and nonstandard labels for linear programs, although we will often resort to standard programs for simplicity.

3. Value Function

Given a primal, consider its value as a function of the right-hand side constraints:

$$V(b) = \sup_{x \geq 0} \{cx : Ax \leq b\}.$$

The function $V : \mathbb{R}^m \to \mathbb{R} \cup \{\pm\infty\}$ describes the value of the primal as the vector b varies. In this section, we will derive the main properties of this value function. Let dom V denote the *effective domain* of V, defined as dom $V = \{b \in \mathbb{R}^m : V(b) > -\infty\}$. The effective domain collects all the right-hand side constraints b of the primal that lead to a nonempty constraint set, i.e., for which there exists $x \geq \mathbf{0}$ such that $Ax \leq b$.

Lemma 7. *$V(\mathbf{0}) = 0$ or $+\infty$, and if $V(\mathbf{0}) = +\infty$ then $V(b) = +\infty$ for all $b \in$ dom V.*

Proof. Clearly, $\mathbf{0} \in \text{dom } V$, since if $b = \mathbf{0}$ then $x = \mathbf{0}$ satisfies $x \geq \mathbf{0}$ and $Ax \leq \mathbf{0}$. Therefore, $V(\mathbf{0}) \geq c \cdot \mathbf{0} = 0$. If $V(\mathbf{0}) > 0$ then there exists $x \geq \mathbf{0}$ such that $Ax \leq \mathbf{0}$ and $cx > 0$. But then λx is also feasible for all $\lambda > 0$, with value $c\lambda x = \lambda(cx)$, which can be made arbitrarily large with λ, since $cx > 0$. Hence, $V(\mathbf{0}) = +\infty$. It remains to show that $V(\mathbf{0}) = +\infty$ implies $V(b) = +\infty$ for all $b \in \text{dom } V$. If $V(\mathbf{0}) = +\infty$, there is a sequence $\{x^n\}$ of feasible vectors, i.e., such that $x^n \geq \mathbf{0}$ and $Ax^n \leq \mathbf{0}$ with $cx^n \to +\infty$ as $n \to \infty$. For any vector $b \in \text{dom } V$, there exists a feasible vector $x \geq \mathbf{0}$ such that $Ax \leq b$. However, the vector $x + x^n \geq \mathbf{0}$ satisfies $A(x+x^n) = Ax + Ax^n \leq b$, so is feasible, and $c(x+x^n) = cx + cx^n \to +\infty$, so $V(b) = +\infty$. □

Lemma 8. *V is positively homogeneous and superadditive, hence concave.*

Proof. Positive homogeneity means $V(\lambda b) = \lambda V(b)$ for every $\lambda > 0$. If $V(b)$ is $\pm\infty$, this is immediate. Next, suppose that $V(b) \in \mathbb{R}$. By Theorem 3 there exists $x^* \geq \mathbf{0}$ such that $Ax^* \leq b$ and $V(b) = cx^*$. For any $\lambda > 0$, clearly $V(\lambda b) \geq c(\lambda x^*)$, since $\lambda x^* \geq \mathbf{0}$ and $A(\lambda x^*) \leq \lambda b$. On the other hand, if $V(\lambda b) > c\lambda x^*$ then there is another vector $x^{**} \geq \mathbf{0}$ such that $Ax^{**} \leq \lambda b$ and $V(\lambda b) = cx^{**} > \lambda cx^*$. But this implies that $z^* = x^{**}/\lambda$ satisfies $z^* \geq \mathbf{0}$, $Az^* \leq b$ and $cz^* > cx^*$, contradicting the optimality of x^*. Therefore, $V(\lambda b) = \lambda V(b)$.

For superadditivity, we must show that $V(b) + V(d) \leq V(b + d)$ for all $b, d \in \mathbb{R}^m$. If either $V(b)$ or $V(d)$ are equal to $-\infty$, we are done. Suppose that both $V(b), V(d) > -\infty$, so there exists $x, z \geq \mathbf{0}$ such that $Ax \leq b$ and $Az \leq d$. Now, letting $w = x + z$, clearly $w \geq \mathbf{0}$ and $Aw = Ax + Az \leq b + d$, therefore, since V is the maximum value among all feasible solutions with right-hand side constraint $b + d$, it follows that $cw \leq V(b + d)$. To obtain concavity, let $b, d \in \mathbb{R}^m$ and $\lambda \in [0, 1]$. By positive homogeneity and superaddivity, clearly $V(\lambda b + (1 - \lambda)d) \geq V(\lambda b) + V((1 - \lambda)d) = \lambda V(b) + (1 - \lambda)V(d)$. □

The *subdifferential* of the concave function V at b is the closed convex set

$$\partial V(b) = \{y \in \mathbb{R}^m : V(b) - yb \geq V(\hat{b}) - y\hat{b} \quad \forall \hat{b} \in \mathbb{R}^m\}.$$

Some people reserve the term subdifferential for convex functions and would call the object above a "superdifferential." We follow Rockafellar and call it the subdifferential of a concave function. At any point of differentiability, the subdifferential of V equals the derivative. If the function V has a kink at some point, the subdifferential is the set of supporting hyperplanes of the function at that point. Another way to think about it in this context is as follows.

Proposition 9. *If $b \in \mathbb{R}^m$ is a point where $V(b)$ is finite then*

$$\partial V(b) = \arg\min_{y \geq 0} \{yb : yA \geq c\}.$$

Proof. Since there is a point $b \in \mathbb{R}^m$ where $V(b)$ is finite, there is strong duality there. Therefore, the dual has a feasible solution, so there is strong duality for all $\hat{b} \in \mathbb{R}^m$. If y solves the dual problem then $V(b) - yb = 0$ by strong duality, but $V(\hat{b}) \leq y\hat{b}$ for all $\hat{b} \in \mathbb{R}^m$ by weak duality, so $V(b) - yb \geq V(\hat{b}) - y\hat{b}$, therefore $y \in \partial V(b)$.

Conversely, suppose that y is not a dual solution. If $yb > V(b)$ then $V(b) - yb < 0 = V(\mathbf{0}) - y\mathbf{0}$, hence $y \notin \partial V(b)$. Now suppose that $yb \leq V(b)$. For y to not be a dual solution, it must be infeasible. Suppose first that this is because $y_j < 0$ for some j. Letting $\mathbf{1}_j$ equal 1 at the jth place and 0 elsewhere, the subdifferential inequality

$$V(b) - yb \geq V(b + \mathbf{1}_j) - y(b + \mathbf{1}_j)$$

implies $0 > y_j \geq V(b+\mathbf{1}_j) - V(b) \geq 0$, where the last inequality follows by revealed preference, since $b \leq b + \mathbf{1}_j$. This contradiction implies that $y \notin \partial V(b)$. Next, suppose that $y \geq \mathbf{0}$ but $yA_i < c_i$ for some i, where A_i is the ith column of A. The subdifferential inequality

$$V(b) - yb \geq V(b + A_i) - y(b + A_i)$$

implies $c_i > yA_i \geq V(b + A_i) - V(b)$. Since $V(b)$ is finite, the dual with objective b has an optimal solution, call it y^0. If $V(b+A_i)$ is finite, there is an optimal dual solution for objective $b + A_i$, call it y^i. By revealed preference, $V(b) = y^0 b \leq y^i b$, hence $c_i > y^i(b + A_i) - y^i b = y^i A_i$, since $y^i(b + A_i) = V(b + A_i)$. This contradicts feasibility of y^i, therefore, $y \notin \partial V(b)$. Finally, suppose that $V(b+A_i)$ is not finite. Since the dual is feasible for all $\hat{b} \in \mathbb{R}^m$, there is a sequence $\{y^{i,n}\}_n$ of feasible dual variables such that $y^{i,n}(b + A_i) \to -\infty$ as $n \to \infty$. In particular, there exists N such that $c_i > y^{i,n}(b+A_i) - V(b)$ for all $n \geq N$. By revealed preference, $V(b) \leq y^{i,n} b$, hence $c_i > y^{i,n}(b + A_i) - y^{i,n} b = y^{i,n} A_i$, contradicting feasibility of $y^{i,n}$, thus $y \notin \partial V(b)$. □