Lecture 3: Linear programming

Lecture 2 develops separation for convex sets and cones. Lecture 3 shows how that geometry specializes to the most classical optimization model: linear programming. In this setting one can see, within a single family of problems, three kinds of certificates at once: dual lower-bound certificates, infeasibility certificates, and optimality certificates.

Example 3.1 (A graduate-student workload LP)**.** Let x_1 denote the number of research-focused work hours this week, and let x_2 denote the number of course-focused work hours. Suppose the week must deliver at least 6 units of research progress and 9 units of course preparation. The per-hour contributions are:

Work type	Research progress (units/hour)	Course preparation (units/hour)
Research-focused work	2	1
Course-focused work	1	2

If the corresponding strain costs are 5 and 4, respectively, then the resulting linear program is

$$\min \left\{5x_1 + 4x_2 : 2x_1 + x_2 \geq 6,\ x_1 + 2x_2 \geq 9,\ x_1 \geq 0,\ x_2 \geq 0\right\}.$$

It turns out the optimal solution is $(x_1, x_2) = (1, 4)$, which is feasible and has objective value

$$5 \cdot 1 + 4 \cdot 4 = 21.$$

To show directly that no feasible point can do better, take a positive linear combination of the two requirement inequalities:

$$2(2x_1 + x_2 \geq 6) + (x_1 + 2x_2 \geq 9).$$

This gives

$$5x_1 + 4x_2 \geq 21.$$

Thus every feasible point has objective value at least 21, and since $(1, 4)$ attains this bound, it is optimal. The weights in this argument are already hinting at the dual problem: the first requirement carries weight 2, while the second carries weight 1.

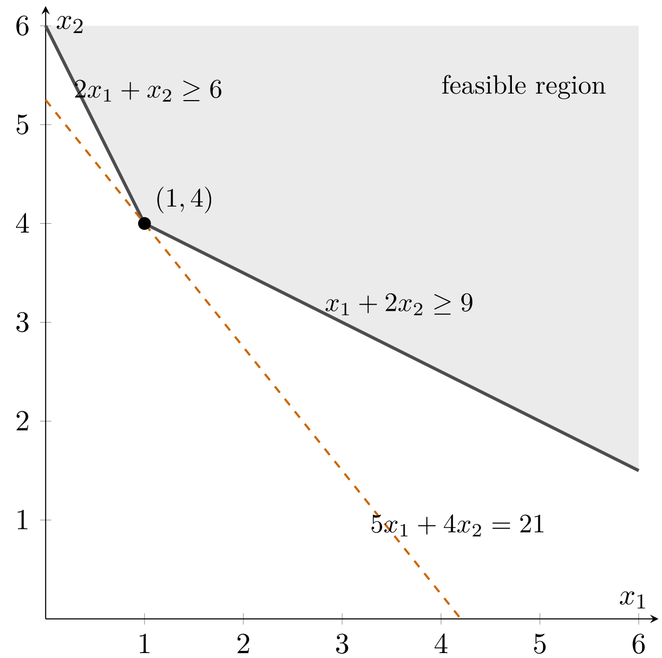


Figure 1: Visible portion of the feasible region in Example 3.1. The dashed line is the objective level set $5x_1 + 4x_2 = 21$, which first meets the feasible region at $(1, 4)$.

Definition 3.1 (Canonical primal and dual LP)**.** Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$. The canonical primal LP is

$$\inf \left\{ c^\top x : Ax \geq b,\ x \geq 0 \right\}.$$

Its dual is

$$\sup \left\{ b^\top y : A^\top y \leq c,\ y \geq 0 \right\}.$$

A vector $x \in \mathbb{R}^n$ satisfying $Ax \geq b$ and $x \geq 0$ is called primal feasible. A vector $y \in \mathbb{R}^m$ satisfying $A^\top y \leq c$ and $y \geq 0$ is called dual feasible.

Lemma 3.1 (Weak duality for LP)**.** *Let $x \in \mathbb{R}^n$ satisfy $Ax \geq b$ and $x \geq 0$, and let $y \in \mathbb{R}^m$ satisfy $A^\top y \leq c$ and $y \geq 0$. Then*

$$b^\top y \leq c^\top x.$$

Lemma 3.1 says that every dual-feasible vector y gives a valid lower bound on the primal objective. In the graduate-student example, the positive linear combination certificate is exactly of this kind: dual feasibility guarantees that the weighted requirements never exceed the objective coefficients, and the resulting right-hand side gives a global lower bound on every primal-feasible point.

> *Proof idea of Lemma 3.1.* The argument is only one line: $Ax \geq b$ and $y \geq 0$ give $b^\top y \leq y^\top Ax$, while $A^\top y \leq c$ and $x \geq 0$ give $y^\top Ax = x^\top A^\top y \leq c^\top x$.

Before proving strong duality, it is useful to isolate one structural LP fact: the sets of attainable primal and dual objective values are closed subsets of $\mathbb{R}$. This is what will turn equality of optimal values into actual attainment of optimal points.

Lemma 3.2 (Objective-value sets are closed). *Define*

$$P := \left\{c^\top x : Ax \geq b,\ x \geq 0\right\}, \qquad D := \left\{b^\top y : A^\top y \leq c,\ y \geq 0\right\}.$$

Then P and D are closed subsets of $\mathbb{R}$.

Theorem 3.3 (Strong duality for LP). *Assume that the primal LP is feasible and has finite optimal value. Then the dual LP is feasible, has finite optimal value, and*

$$\inf\left\{c^\top x : Ax \geq b,\ x \geq 0\right\} = \sup\left\{b^\top y : A^\top y \leq c,\ y \geq 0\right\}.$$

Moreover, there exist $x^\star \in \mathbb{R}^n$ and $y^\star \in \mathbb{R}^m$ such that

$$Ax^\star \geq b, \qquad x^\star \geq 0, \qquad A^\top y^\star \leq c, \qquad y^\star \geq 0,$$

and

$$c^\top x^\star = b^\top y^\star = \inf\left\{c^\top x : Ax \geq b,\ x \geq 0\right\} = \sup\left\{b^\top y : A^\top y \leq c,\ y \geq 0\right\}.$$

Theorem 3.4 (Farkas' lemma). *Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Exactly one of the following two statements holds:*

1. *there exists $x \in \mathbb{R}^n$ such that $Ax \geq b$ and $x \geq 0$;*
2. *there exists $y \in \mathbb{R}^m$ such that $A^\top y \leq 0$, $y \geq 0$, and $b^\top y > 0$.*

Theorem 3.4 is the main theorem in this lecture for certifying infeasibility. It says that if the system $Ax \geq b$, $x \geq 0$ has no solution, then there is a nonnegative vector y such that $A^\top y \leq 0$ and $b^\top y > 0$. One can view this as a contradiction witness for the requirement inequalities together with the sign restriction $x \geq 0$: any feasible x would force

$$b^\top y \leq y^\top Ax = x^\top A^\top y \leq 0,$$

so such a vector y explicitly rules out feasibility.

Proof idea of Theorem 3.4. Let $K := -A\mathbb{R}^n_+ + \mathbb{R}^m_+$. Then $Ax \geq b$ with $x \geq 0$ is equivalent to the membership statement $-b \in K$: the slack vector $s := Ax - b$ is nonnegative, so $-b = -Ax + s$. Thus infeasibility means exactly $-b \notin K$. Now

$$K = \operatorname{cone}\left\{-Ae_1, \ldots, -Ae_n, e_1, \ldots, e_m\right\},$$

so Lemma 3.7 shows that K is closed. Therefore Theorem 2.7 gives a separating covector y with $\langle y, z\rangle \geq 0$ for every $z \in K$ and $\langle y, -b\rangle < 0$. Testing this inequality on vectors of the form $z = s$ with $s \geq 0$ forces $y \geq 0$, and testing it on vectors of the form $z = -Ax$ with $x \geq 0$ forces $A^\top y \leq 0$. These are exactly the coordinate conditions in the certificate.

Theorem 3.3 answers the main question left open by Lemma 3.1: how tight can a dual lower bound be? It says that, for linear programs, the best dual certificate is exactly tight. In fact, both the primal and the dual optima are attained.

Proof idea of Theorem 3.3. Now use Theorem 3.4 as an infeasibility-certificate theorem. Look below the true optimum: if one asks for a value strictly smaller than the primal optimum, then the augmented system

$$Ax \geq b, \qquad x \geq 0, \qquad c^\top x \leq \mu$$

becomes infeasible. Applying Theorem 3.4 to this enlarged system produces a certificate (y, α), and the key step is to show that $\alpha > 0$. After normalizing by α, one gets a dual-feasible point with value strictly above μ. Letting $\mu \uparrow p^\star$ forces the dual value to reach the primal optimum. Finally, Lemma 3.2 shows that the primal and dual objective-value sets are closed subsets of $\mathbb{R}$. Therefore $p^\star$ belongs to both objective-value sets, so both optima are attained.

Theorem 3.5 (Complementary slackness)**.** *Let $x^\star \in \mathbb{R}^n$ and $y^\star \in \mathbb{R}^m$. The following are equivalent:*

1. *$x^\star$ is primal feasible, $y^\star$ is dual feasible, and both are optimal;*

2. *$x^\star$ is primal feasible, $y^\star$ is dual feasible, and*

$$\forall i \in \{1, \ldots, m\}, \qquad y_i^\star((Ax^\star)_i - b_i) = 0,$$

and

$$\forall j \in \{1, \ldots, n\}, \qquad x_j^\star(c_j - (A^\top y^\star)_j) = 0.$$

Finally, Theorem 3.5 characterizes exactly when equality holds in weak duality. At optimality, every positive dual multiplier must sit on a tight primal requirement, and every positive primal variable must sit on a tight dual inequality. This is the first instance of a pattern that will recur throughout the course: optimality is not only about matching objective values, but also about matching the geometry of the active constraints to the geometry of the dual certificate.

Proof idea of Theorem 3.5. Expand the duality gap as

$$c^\top x^\star - b^\top y^\star = x^{\star\top}(c - A^\top y^\star) + y^{\star\top}(Ax^\star - b),$$

a sum of nonnegative terms. Equality of objective values therefore forces every term to vanish, and that vanishing is exactly the complementary-slackness condition.

3.1 Value functions and shadow prices

Weak duality gives a lower bound on the optimal value of the single LP with right-hand side b. But one can understand the geometry of the dual variable much better by enlarging the picture: instead of fixing b, ask how the optimal cost changes when the requirement vector is perturbed. In this lecture, a shadow price means exactly such a sensitivity coefficient: it records how the optimal value changes when one requirement is tightened. This is the first place where shadow prices begin to look like sensitivity information rather than merely algebraic multipliers.

Definition 3.2 (Requirement-perturbation value function for an LP)**.** Fix $A \in \mathbb{R}^{m \times n}$ and $c \in \mathbb{R}^n$. The requirement-perturbation value function associated with the canonical LP is

$$V : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}, \qquad V(u) := \inf\left\{c^\top x : Ax \geq u,\ x \geq 0\right\}.$$

Lemma 3.6 (The LP value function is convex)**.** *The function V in Definition 3.2 is convex.*

Now weak duality applies not only to the original LP, but simultaneously to the whole perturbed family. If $y \in \mathbb{R}^m$ satisfies $A^\top y \le c$ and $y \ge 0$, then the same calculation as in Lemma 3.1 shows that

$$V(u) \ge u^\top y.$$

Thus every dual-feasible vector y gives an affine lower bound on the entire value function V, not just on the single number $V(b)$. In this enlarged picture, the meaning of the dual variable becomes much easier to see: it measures how the optimal cost responds to changes in the requirement vector, which is exactly the usual economic interpretation of a shadow price. If strong duality holds at a chosen base requirement vector b and $y^\star$ is dual optimal there, then

$$V(b) = b^\top y^\star.$$

So an optimal dual point is not just an algebraic multiplier: it is a supporting covector of the convex value function V at the base point b. In particular, if V happens to be differentiable at b, then this supporting covector is unique and

$$y^\star = \nabla V(b).$$

This is the cleanest economic interpretation of a shadow price: the component $y_i^\star$ is the marginal increase in optimal cost per unit increase in the i-th requirement, at least to first order near the base point b. When V is not differentiable, one should replace the gradient by a supporting covector, equivalently a subgradient. Lecture 4 abstracts exactly this LP picture into the general marginal value function $p(u) = \inf_x \Phi(x, u)$.

Example 3.2 (The workload value function and shadow prices)**.** Return to Example 3.1, but now replace the fixed requirement vector $(6, 9)$ by a variable pair $(u_1, u_2) \in \mathbb{R}^2$. The resulting value function is

$$V(u_1, u_2) := \inf \{5x_1 + 4x_2 : 2x_1 + x_2 \ge u_1,\ x_1 + 2x_2 \ge u_2,\ x_1 \ge 0,\ x_2 \ge 0\}.$$

The dual feasible region is the polygon

$$\Big\{(p_1, p_2) \in \mathbb{R}^2 : 2p_1 + p_2 \le 5,\ p_1 + 2p_2 \le 4,\ p_1 \ge 0,\ p_2 \ge 0\Big\},$$

whose vertices are

$$(0,0), \qquad \left(\frac{5}{2}, 0\right), \qquad (2,1), \qquad (0,2).$$

By LP strong duality, $V(u_1, u_2)$ is therefore the pointwise maximum of the four affine functions

$$0, \qquad \frac{5}{2}u_1, \qquad 2u_1 + u_2, \qquad 2u_2 :$$

$$V(u_1, u_2) = \max\left\{0,\ \frac{5}{2}u_1,\ 2u_1 + u_2,\ 2u_2\right\}.$$

At the base point $(u_1, u_2) = (6, 9)$, the active piece is

$$2u_1 + u_2,$$

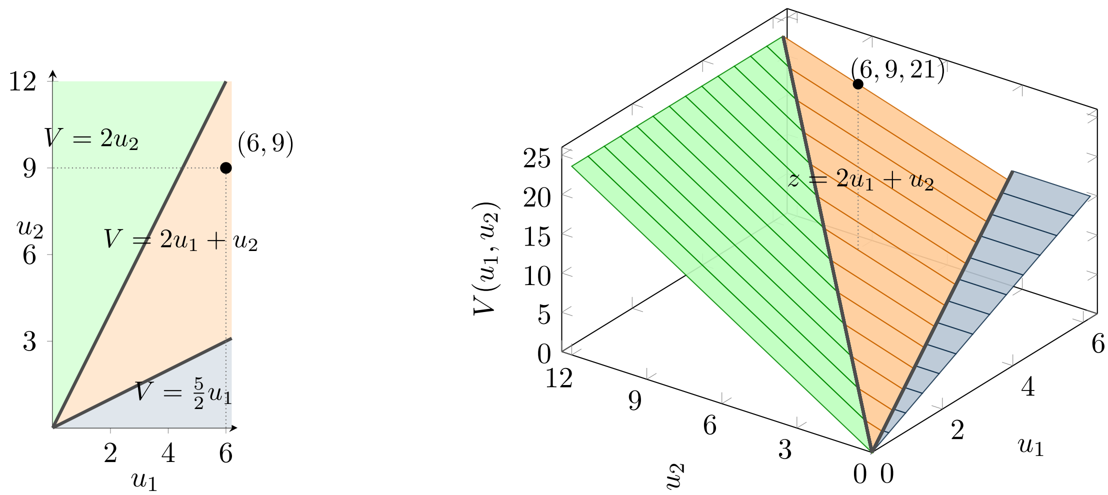
so

$$V(6, 9) = 2 \cdot 6 + 9 = 21.$$

In this example, V is differentiable at $(6, 9)$, and

$$\nabla V(6, 9) = (2, 1).$$

So near $(6, 9)$, increasing the first requirement by one unit raises the optimal cost by about 2, while increasing the second requirement by one unit raises it by about 1. This is the easiest case to interpret economically: the shadow price is literally the gradient of the value function. In general, however, LP value functions are only piecewise linear and need not be differentiable at kinks; then the correct general object is not a unique gradient, but a supporting covector, equivalently a subgradient.



Active affine piece in the (u_1, u_2)-plane

The highlighted face gives the shadow price $(2, 1)$

Figure 2: The value function for Example 3.1, namely $V(u_1, u_2) = \max\{0, \frac{5}{2}u_1, 2u_1 + u_2, 2u_2\}$. The left panel shows which affine piece is active at each requirement vector (u_1, u_2); the right panel shows the corresponding polyhedral graph. At the base point $(6, 9)$, the active face is $z = 2u_1 + u_2$, so the shadow price is $(2, 1)$.

Dependency and proof sketch

1. Lemma 3.1 is one line:
$$b^\top y \le y^\top A x = x^\top A^\top y \le x^\top c = c^\top x$$
because $y \ge 0$, $Ax \ge b$, $x \ge 0$, and $A^\top y \le c$.

2. Theorem 3.3 is the first clean specialization of duality to a concrete family of problems. It says that the best lower bound produced by Lemma 3.1 is exactly tight. The proof below derives it from Theorem 3.4 by applying Farkas to the augmented infeasible system
$$Ax \ge b, \qquad x \ge 0, \qquad c^\top x \le \mu.$$

3. Theorem 3.4 is the infeasibility theorem for linear inequalities. Its visible content is simple: if the system $Ax \ge b$, $x \ge 0$ has no solution, then there is a vector witness ruling it out. The proof uses the cone $-A\mathbb{R}^n_+ + \mathbb{R}^m_+$ together with the cone-separation theorem from Lecture 2 to produce exactly such a witness $y \ge 0$, $A^\top y \le 0$, $b^\top y > 0$.

4. Theorem 3.5 is weak duality plus the observation that the duality gap is exactly the complementary-slackness sum

$$c^\top x^\star - b^\top y^\star = \sum_{i=1}^{m} y_i^\star\big((Ax^\star)_i - b_i\big) \;+\; \sum_{j=1}^{n} x_j^\star\big(c_j - (A^\top y^\star)_j\big).$$

5. Lemma 3.6 is a direct convexification argument: feasible points for right-hand sides u^1 and u^2 can be averaged, so the intermediate right-hand side inherits a feasible point with the corresponding convex-combination objective value.

Proofs

Proof of Lemma 3.6. First proof: direct from the definition of convexity. Fix $u^1, u^2 \in \mathbb{R}^m$ and $\theta \in [0,1]$. If either $V(u^1)$ or $V(u^2)$ is $+\infty$, then the convexity inequality is immediate, so assume both are finite. Let $\varepsilon > 0$. Choose $x^1, x^2 \in \mathbb{R}^n$ such that

$$Ax^1 \ge u^1, \quad x^1 \ge 0, \quad c^\top x^1 \le V(u^1) + \varepsilon,$$

and

$$Ax^2 \ge u^2, \quad x^2 \ge 0, \quad c^\top x^2 \le V(u^2) + \varepsilon.$$

Then

$$\bar{x} := \theta x^1 + (1-\theta)x^2$$

satisfies $\bar{x} \ge 0$ and

$$A\bar{x} = \theta A x^1 + (1-\theta)Ax^2 \ge \theta u^1 + (1-\theta)u^2.$$

Hence $\bar{x}$ is feasible for the right-hand side $\theta u^1 + (1-\theta)u^2$. Therefore

$$V(\theta u^1 + (1-\theta)u^2) \le c^\top \bar{x} = \theta c^\top x^1 + (1-\theta)c^\top x^2 \le \theta V(u^1) + (1-\theta)V(u^2) + \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, it follows that

$$V(\theta u^1 + (1-\theta)u^2) \le \theta V(u^1) + (1-\theta)V(u^2).$$

Geometric viewpoint (after the projection exercise at the end of this lecture). Consider the set

$$\mathcal{E} := \Big\{(u,x,t) \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R} : Ax \ge u,\ x \ge 0,\ c^\top x \le t\Big\}.$$

This is a convex set: each defining constraint is affine in (u,x,t), so $\mathcal{E}$ is an intersection of affine halfspaces. By definition,

$$(u,t) \in \pi_{u,t}(\mathcal{E}) \quad \iff \quad \exists x \ge 0 \text{ such that } Ax \ge u \text{ and } c^\top x \le t,$$

which is exactly the condition $V(u) \le t$. Hence

$$\operatorname{epi}(V) = \pi_{u,t}(\mathcal{E}).$$

In other words, taking the partial infimum over x is exactly the same as projecting the epigraph onto the (u,t)-coordinates. By the projection exercise at the end of this lecture, projections preserve convexity. Therefore $\operatorname{epi}(V)$ is convex, and thus V is convex. □

Proof of Lemma 3.1. Because $y \geq 0$ and $Ax \geq b$, one has

$$b^\top y \leq y^\top Ax.$$

Transposing the scalar gives

$$y^\top Ax = x^\top A^\top y.$$

Since $A^\top y \leq c$ coordinatewise and $x \geq 0$ coordinatewise,

$$x^\top A^\top y \leq x^\top c.$$

Therefore

$$b^\top y \leq c^\top x.$$

□

Proof of Theorem 3.4. We first show that the two statements cannot hold simultaneously. Assume that there exist $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ such that

$$Ax \geq b, \quad x \geq 0, \quad A^\top y \leq 0, \quad y \geq 0, \quad b^\top y > 0.$$

Then

$$b^\top y \leq y^\top Ax = x^\top A^\top y \leq 0,$$

which contradicts $b^\top y > 0$. Thus at most one statement can hold.
We now show that at least one statement holds. Define

$$K := -A\mathbb{R}^n_+ + \mathbb{R}^m_+ = \left\{-Ax + s : x \in \mathbb{R}^n_+, \ s \in \mathbb{R}^m_+\right\}.$$

Since

$$K = \text{cone}\left\{-Ae_1, \dots, -Ae_n, e_1, \dots, e_m\right\},$$

Lemma 3.7 shows that K is a closed convex cone. Statement (1) is equivalent to $b \in A\mathbb{R}^n_+ - \mathbb{R}^m_+$, hence to $-b \in K$. If statement (1) fails, then $-b \notin K$. Applying the cone separation theorem Theorem 2.7 to the closed convex cone K and the point $-b \notin K$, we obtain a vector $y \in \mathbb{R}^m$ such that

$$\forall z \in K, \qquad \langle y, z \rangle \geq 0, \qquad \text{and} \qquad \langle y, -b \rangle < 0.$$

For each standard basis vector $e_i \in \mathbb{R}^m$, one has $e_i \in \mathbb{R}^m_+ \subseteq K$, and therefore

$$\langle y, e_i \rangle = y_i \geq 0.$$

Thus $y \geq 0$. Moreover, for every $x \in \mathbb{R}^n_+$, one has $-Ax \in K$, so

$$\langle y, -Ax \rangle \geq 0.$$

Equivalently,

$$x^\top A^\top y \leq 0 \qquad \forall x \in \mathbb{R}^n_+.$$

Taking $x = e_j$ shows $(A^\top y)_j \leq 0$ for every j, so $A^\top y \leq 0$. Finally,

$$b^\top y = -\langle y, -b \rangle > 0.$$

Thus statement (2) holds. This proves that exactly one of the two statements is true. □

Lemma 3.7 (Finitely generated cones are closed)**.** *For vectors $v_1, \dots, v_N \in \mathbb{R}^m$, the cone*

$$\operatorname{cone}\{v_1, \dots, v_N\} := \left\{ \sum_{i=1}^{N} \lambda_i v_i : \lambda_i \geq 0 \text{ for all } i \right\}$$

is a closed convex cone.

Proof of Lemma 3.7. Convexity is immediate from the definition, so we only prove closedness. Let $K := \operatorname{cone}\{v_1, \dots, v_N\}$, and suppose $x_k \in K$ with $x_k \to x$. For each k, choose a nonnegative representation

$$x_k = \sum_{i=1}^{N} \lambda_{k,i} v_i, \qquad \lambda_{k,i} \geq 0,$$

whose support $\{i : \lambda_{k,i} > 0\}$ is minimal. Then the active generators are linearly independent: otherwise a nontrivial linear relation among them would let us perturb the coefficients, preserve the same point, keep all coefficients nonnegative, and eliminate at least one positive coefficient, contradicting minimality.

There are only finitely many possible supports, so after passing to a subsequence we may assume that all x_k use the same active support $I \subseteq \{1, \dots, N\}$. Writing

$$x_k = \sum_{i \in I} \lambda_{k,i} v_i,$$

the coefficient map

$$T_I : \mathbb{R}^I \to \operatorname{span}\{v_i : i \in I\}, \qquad (\lambda_i)_{i \in I} \mapsto \sum_{i \in I} \lambda_i v_i$$

is a linear isomorphism because $\{v_i\}_{i \in I}$ is linearly independent. Hence T_I^{-1} is continuous, so the coefficient vectors $(\lambda_{k,i})_{i \in I} = T_I^{-1}(x_k)$ converge to some $(\lambda_i)_{i \in I} \in \mathbb{R}^I$. Since each $\lambda_{k,i} \geq 0$, also $\lambda_i \geq 0$ for every $i \in I$. Passing to the limit gives

$$x = \sum_{i \in I} \lambda_i v_i \in K.$$

Thus K is closed. □

Proof of Lemma 3.2. Let

$$P := \left\{ c^\top x : Ax \geq b,\ x \geq 0 \right\}, \qquad D := \left\{ b^\top y : A^\top y \leq c,\ y \geq 0 \right\}.$$

For $r \in \mathbb{R}$, one has $r \in P$ if and only if there exist $x \geq 0$ and $s \geq 0$ such that

$$\begin{pmatrix} A & -I_m \\ c^\top & 0 \end{pmatrix} \begin{pmatrix} x \\ s \end{pmatrix} = \begin{pmatrix} b \\ r \end{pmatrix},$$

so P is the affine slice $\{r \in \mathbb{R} : (b, r) \in K_P\}$ of the finitely generated cone

$$K_P := \operatorname{cone}\left(\text{columns of } \begin{pmatrix} A & -I_m \\ c^\top & 0 \end{pmatrix} \right).$$

By Lemma 3.7, the cone K_P is closed. Therefore P is closed. Similarly, $r \in D$ if and only if there exist $y \geq 0$ and $s \geq 0$ such that

$$\begin{pmatrix} A^\top & I_m \\ b^\top & 0 \end{pmatrix} \begin{pmatrix} y \\ s \end{pmatrix} = \begin{pmatrix} c \\ r \end{pmatrix},$$

so D is the affine slice $\{r \in \mathbb{R} : (c, r) \in K_D\}$ of the finitely generated cone

$$K_D := \operatorname{cone}\left(\text{columns of } \begin{pmatrix} A^\top & I_m \\ b^\top & 0 \end{pmatrix} \right).$$

Again by Lemma 3.7, the cone K_D is closed. Hence D is closed as well. □

Proof of Theorem 3.3.

$$p^* := \inf \left\{ c^\top x : Ax \geq b,\ x \geq 0 \right\}.$$

By hypothesis, $p^* \in \mathbb{R}$ and the primal feasible set is nonempty. Fix any scalar $\mu < p^*$. Then the system

$$Ax \geq b, \qquad x \geq 0, \qquad c^\top x \leq \mu$$

is infeasible. Rewrite this as

$$\begin{pmatrix} A \\ -c^\top \end{pmatrix} x \geq \begin{pmatrix} b \\ -\mu \end{pmatrix}.$$

By Theorem 3.4, there exists $(y, \alpha) \in \mathbb{R}^m \times \mathbb{R}$ such that

$$\begin{pmatrix} A \\ -c^\top \end{pmatrix}^\top \begin{pmatrix} y \\ \alpha \end{pmatrix} \leq 0, \qquad \begin{pmatrix} y \\ \alpha \end{pmatrix} \geq 0, \qquad \begin{pmatrix} b \\ -\mu \end{pmatrix}^\top \begin{pmatrix} y \\ \alpha \end{pmatrix} > 0.$$

Equivalently,

$$A^\top y - \alpha c \leq 0, \qquad y \geq 0, \qquad \alpha \geq 0, \qquad b^\top y - \alpha\mu > 0.$$

We claim that $\alpha > 0$. Indeed, if $\alpha = 0$, then $A^\top y \leq 0$, $y \geq 0$, and $b^\top y > 0$, which contradicts primal feasibility by Theorem 3.4. Thus $\alpha > 0$. Define

$$\bar{y} := \frac{y}{\alpha}.$$

Then

$$A^\top \bar{y} \leq c, \qquad \bar{y} \geq 0,$$

so $\bar{y}$ is dual feasible, and

$$b^\top \bar{y} = \frac{b^\top y}{\alpha} > \mu.$$

Because $\mu < p^*$ was arbitrary, this shows that for every $\mu < p^*$ there exists a dual-feasible point with objective value strictly larger than μ. Therefore the dual optimal value d^* satisfies

$$d^* \geq p^*.$$

On the other hand, by Lemma 3.1, every dual-feasible value is at most every primal-feasible value, so

$$d^* \leq p^*.$$

Hence

$$p^* = d^*.$$

Let

$$P := \left\{c^\top x : Ax \geq b,\ x \geq 0\right\}, \qquad D := \left\{b^\top y : A^\top y \leq c,\ y \geq 0\right\}.$$

By primal feasibility, P is nonempty and $p^* = \inf P$. By Lemma 3.2, the set P is closed. Since $P \subseteq \mathbb{R}$ is nonempty and bounded below, its infimum belongs to P. Hence there exists a primal-feasible point $x^\star$ such that

$$c^\top x^\star = p^*.$$

Next, the argument above shows that for every $\mu < p^*$ there exists a dual-feasible point with objective value strictly larger than μ. Thus D is nonempty and

$$\sup D = p^*.$$

By Lemma 3.1, every dual-feasible value is a lower bound on the primal objective, so

$$D \subseteq (-\infty, p^*].$$

Again by Lemma 3.2, the set D is closed. Since $D \subseteq \mathbb{R}$ is nonempty and bounded above, its supremum belongs to D. Therefore there exists a dual-feasible point $y^\star$ such that

$$A^\top y^\star \leq c, \qquad y^\star \geq 0, \qquad b^\top y^\star = \sup D = p^* = d^*.$$

Thus both the primal and dual optima are attained. □

Proof of Theorem 3.5. Assume first that item (1) holds. Then $x^\star$ is primal feasible, $y^\star$ is dual feasible, and both are optimal. By Lemma 3.1,

$$b^\top y^\star \leq c^\top x^\star.$$

Because both points are optimal, equality must hold:

$$c^\top x^\star = b^\top y^\star.$$

Using dual feasibility,

$$c^\top x^\star - b^\top y^\star = x^{\star\top}(c - A^\top y^\star) + y^{\star\top}(Ax^\star - b) = \sum_{j=1}^{n} x_j^\star(c_j - (A^\top y^\star)_j) \ + \ \sum_{i=1}^{m} y_i^\star((Ax^\star)_i - b_i).$$

Each term in the sum is nonnegative because $x^\star \geq 0$, $c - A^\top y^\star \geq 0$, $y^\star \geq 0$, and $Ax^\star - b \geq 0$. Since the sum equals zero, every term must be zero. Thus item (2) holds.
Assume now that item (2) holds. Then $x^\star$ is primal feasible, $y^\star$ is dual feasible, and

$$\forall i \in \{1, \dots, m\}, \qquad y_i^\star((Ax^\star)_i - b_i) = 0,$$

and

$$\forall j \in \{1, \dots, n\}, \qquad x_j^\star(c_j - (A^\top y^\star)_j) = 0.$$

Therefore

$$c^\top x^\star - b^\top y^\star = x^{\star\top}(c - A^\top y^\star) + y^{\star\top}(Ax^\star - b) = 0.$$

Hence

$$b^\top y^\star = c^\top x^\star.$$

By Lemma 3.1, every dual-feasible objective value is at most every primal-feasible objective value. Since equality is achieved by $(x^\star, y^\star)$, both points are optimal. Thus item (1) holds. □

Exercises

1. Translate each of the following LP variants into the canonical model of Definition 3.1, and then derive the dual directly in the natural variables of the original formulation.

   (a) The resource-allocation LP

   $$\max\left\{d^\top z : Bz \le h,\ z \ge 0\right\}.$$

   After deriving its dual, interpret the dual variables as prices for the resource constraints.

   (b) The equality-constrained LP

   $$\min\left\{c^\top x : Ax = b,\ x \ge 0\right\}.$$

   Explain why the equality constraint in the primal produces a dual variable with no sign restriction.

   (c) The free-variable LP

   $$\min\left\{c^\top x : Ax \ge b,\ x \in \mathbb{R}^n\right\}.$$

   Write each free variable as a difference of two nonnegative variables, and explain why the absence of a sign restriction on x turns the dual inequality $A^\top y \le c$ into an equality.

   (d) The bounded-variable LP

   $$\min\left\{c^\top x : Ax \ge b,\ 0 \le x \le u\right\}.$$

   Derive its dual in a form that keeps separate the dual variables for the requirement constraints $Ax \ge b$ and for the upper bounds $x \le u$.

2. Let $C \subseteq \mathbb{R}^{m+n}$ and $K, L \subseteq \mathbb{R}^n$ be convex sets, and let $T : \mathbb{R}^n \to \mathbb{R}^p$ be an affine map.

   (a) Show that the coordinate projection

   $$\pi(C) := \{u \in \mathbb{R}^m : \exists x \in \mathbb{R}^n \text{ such that } (u, x) \in C\}$$

   is convex.

   (b) Show that the Minkowski sum

   $$K + L := \{x + y : x \in K,\ y \in L\}$$

   is convex.

   (c) Show that the affine image $T(K)$ is convex.

   (d) Show that the affine preimage

   $$T^{-1}(L) := \{x \in \mathbb{R}^n : T(x) \in L\}$$

   is convex.

   Which part of Lemma 3.6 can be reinterpreted using part (1)?