Linear Programming - Derivations & Proofs

Paul F. Roysdon, Ph.D.

Contents

1 Mathematical Derivations & Proofs

1.1 Introduction

Linear Programming (LP) optimizes a linear objective subject to linear constraints. The feasible set is a (convex) polyhedron and every LP attains its optimum at an extreme point (vertex) whenever an optimum exists. We derive LP from first principles: standard form, geometry (extreme points), *simplex* optimality and pivot rules, Lagrangian duality (weak/strong duality and complementary slackness), the KKT system, and a primal–dual *interior-point* (barrier) method. All variables and dimensions adhere to the notation below.

1.2 Data and Notation

Let

$$\mathbf{A} \in \mathbb{R}^{m\times n}, \qquad \mathbf{b} \in \mathbb{R}^{m}, \qquad \mathbf{c} \in \mathbb{R}^{n}.$$

Decision variables $\mathbf{x} \in \mathbb{R}^n$ are *columns*. We use the *standard form* LP:

$$\min_{\mathbf{x}\in\mathbb{R}^n} \ \mathbf{c}^\top\mathbf{x} \quad \text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0}. \tag{1}$$

Here $\mathbf{x} \geq \mathbf{0}$ is componentwise. The feasible region is $\mathcal{P} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}\}$.

Dimensions & properties: m constraints, n variables; typically $m \leq n$ for basic solutions to exist. We assume $\mathbf{A}$ has rank $r \leq m$ (full-row rank $r = m$ in the nondegenerate case).

1.3 Model Formulation and Equivalent Forms

General forms with inequalities and free variables can be converted into (1) by adding *slack* variables for "$\leq$" constraints, negating "$\geq$" constraints, and splitting free variables $x_j = x_j^+ - x_j^-$ with $x_j^\pm \geq 0$.

1.4 Geometry: Extreme Points and the Fundamental Theorem

The feasible set $\mathcal{P}$ is a convex polyhedron. A point $\mathbf{x} \in \mathcal{P}$ is an *extreme point* (vertex) if it cannot be written as a nontrivial convex combination of two distinct feasible points.

Basic (feasible) solutions. Let $\mathcal{B} \subset \{1, \ldots, n\}$ be an index set of size m such that the square submatrix $\mathbf{B} \triangleq \mathbf{A}_{(:,\mathcal{B})} \in \mathbb{R}^{m \times m}$ is nonsingular. Partition columns into basic $\mathbf{B}$ and nonbasic $\mathbf{N} = \mathbf{A}_{(:,\mathcal{N})}$ with $\mathcal{N} = \{1, \ldots, n\} \setminus \mathcal{B}$. Define the *basic solution*

$$\mathbf{x}_{\mathcal{N}} = \mathbf{0}, \qquad \mathbf{x}_{\mathcal{B}} = \mathbf{B}^{-1}\mathbf{b}. \tag{2}$$

If additionally $\mathbf{x}_{\mathcal{B}} \geq \mathbf{0}$, then $\mathbf{x}$ is a *basic feasible solution* (BFS).

Extreme points $\iff$ BFS. Every BFS is an extreme point of $\mathcal{P}$; conversely, under $\operatorname{rank}(\mathbf{A}) = m$, every extreme point of $\mathcal{P}$ is a BFS.

Proof. (i) If $\mathbf{x}$ is a BFS with $\mathbf{x}_{\mathcal{N}} = \mathbf{0}$ and $\mathbf{x}_{\mathcal{B}} = \mathbf{B}^{-1}\mathbf{b}$, any feasible $\mathbf{d}$ with $\mathbf{A}\mathbf{d} = \mathbf{0}$ that keeps $\mathbf{x} \pm \epsilon\mathbf{d} \geq \mathbf{0}$ for small ϵ must satisfy $\mathbf{d}_{\mathcal{N}} = \mathbf{0}$ and $\mathbf{d}_{\mathcal{B}} = \mathbf{0}$, so $\mathbf{x}$ cannot be a nontrivial convex combination (extreme). (ii) Conversely, an extreme point must have at least $n - m$ zero components and the active column set be nonsingular, yielding (2). ∎

Fundamental theorem of LP. If (1) has an optimal solution and $\mathcal{P} \neq \emptyset$ and is bounded in the direction of $-\mathbf{c}$, then there exists an optimal BFS (i.e., an optimal extreme point). *Reason:* linear objectives attain extrema over polytopes at extreme points.

1.5 Algebra of the Simplex Method: Optimality and Pivots

At a BFS determined by basis $\mathcal{B}$, write the objective with nonbasic variables $\mathbf{x}_{\mathcal{N}}$ as free:

$$\mathbf{A}_{\mathcal{B}}\mathbf{x}_{\mathcal{B}} + \mathbf{A}_{\mathcal{N}}\mathbf{x}_{\mathcal{N}} = \mathbf{b} \quad \Rightarrow \quad \mathbf{x}_{\mathcal{B}} = \mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{N}\mathbf{x}_{\mathcal{N}}.$$

Let $\mathbf{c}_{\mathcal{B}}$ and $\mathbf{c}_{\mathcal{N}}$ be the corresponding cost partitions. Define the *basic dual* (or simplex multiplier)

$$\mathbf{y}^\top \triangleq \mathbf{c}_{\mathcal{B}}^\top \mathbf{B}^{-1} \in \mathbb{R}^{1 \times m}. \tag{3}$$

The objective as a function of $\mathbf{x}_{\mathcal{N}}$ becomes

$$\begin{aligned} \mathbf{c}^\top \mathbf{x} &= \mathbf{c}_{\mathcal{B}}^\top \mathbf{x}_{\mathcal{B}} + \mathbf{c}_{\mathcal{N}}^\top \mathbf{x}_{\mathcal{N}} = \mathbf{c}_{\mathcal{B}}^\top \left(\mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{N}\mathbf{x}_{\mathcal{N}}\right) + \mathbf{c}_{\mathcal{N}}^\top \mathbf{x}_{\mathcal{N}} \\ &= \underbrace{\mathbf{c}_{\mathcal{B}}^\top \mathbf{B}^{-1}\mathbf{b}}_{\text{current value}} + \underbrace{\left(\mathbf{c}_{\mathcal{N}}^\top - \mathbf{c}_{\mathcal{B}}^\top \mathbf{B}^{-1}\mathbf{N}\right)}_{=\bar{\mathbf{c}}_{\mathcal{N}}^\top} \mathbf{x}_{\mathcal{N}} = \mathbf{y}^\top \mathbf{b} + \bar{\mathbf{c}}_{\mathcal{N}}^\top \mathbf{x}_{\mathcal{N}}. \end{aligned} \tag{4}$$

Here $\bar{\mathbf{c}}_{\mathcal{N}} = \mathbf{c}_{\mathcal{N}} - \mathbf{N}^\top \mathbf{y} \in \mathbb{R}^{n-m}$ are the *reduced costs.*

Simplex optimality condition (primal simplex). At a BFS, if all reduced costs satisfy

$$\bar{\mathbf{c}}_{\mathcal{N}} \;\geq\; \mathbf{0} \quad \text{(componentwise)}, \tag{5}$$

then the BFS is optimal. *Reason:* Any feasible $\mathbf{x}_{\mathcal{N}} \geq \mathbf{0}$ increases the objective by $\bar{\mathbf{c}}_{\mathcal{N}}^\top \mathbf{x}_{\mathcal{N}} \geq 0$.

Improvement direction and ratio test. If some $\bar{c}_q < 0$ (for a *min* problem), variable x_q is a candidate to *enter* the basis. Let the associated column be $\mathbf{a}_q = \mathbf{A}_{(:,q)}$ and direction in basic space

$$\mathbf{d}_{\mathcal{B}} \triangleq -\mathbf{B}^{-1}\mathbf{a}_q, \qquad \mathbf{d}_{\mathcal{N}} = \mathbf{e}_q,$$

so the primal move $\mathbf{x}(\theta) = \mathbf{x} + \theta\,\mathbf{d}$ maintains feasibility $\mathbf{A}\mathbf{x}(\theta) = \mathbf{b}$. To maintain $\mathbf{x}_{\mathcal{B}}(\theta) = \mathbf{x}_{\mathcal{B}} + \theta\mathbf{d}_{\mathcal{B}} \geq \mathbf{0}$, choose

$$\theta^\star = \min_{i:\ (\mathbf{d}_{\mathcal{B}})_i < 0} \frac{(\mathbf{x}_{\mathcal{B}})_i}{-(\mathbf{d}_{\mathcal{B}})_i}. \tag{6}$$

The limiting index p *leaves* the basis; pivot (p, q) yields a new BFS with objective decrease $\Delta z = \theta^\star\,\bar{c}_q < 0$.

Simplex algorithm (high level). Repeat: compute reduced costs; if (5) holds, stop; else pick entering q with $\bar{c}_q < 0$, compute $\theta^\star$ via (6); pivot. Degeneracy (ties with $\theta^\star = 0$) requires anti-cycling rules (e.g., Bland's rule).

1.6 Lagrangian Duality and the Dual LP

Associate multipliers $\mathbf{y} \in \mathbb{R}^m$ for equality constraints and $\mathbf{s} \in \mathbb{R}^n_{\geq 0}$ for $\mathbf{x} \geq \mathbf{0}$ (equivalently, $-\mathbf{x} \leq \mathbf{0}$). The Lagrangian is

$$\mathcal{L}(\mathbf{x}, \mathbf{y}, \mathbf{s}) = \mathbf{c}^\top\mathbf{x} + \mathbf{y}^\top(\mathbf{b} - \mathbf{A}\mathbf{x}) - \mathbf{s}^\top\mathbf{x} = \mathbf{y}^\top\mathbf{b} + (\mathbf{c} - \mathbf{A}^\top\mathbf{y} - \mathbf{s})^\top\mathbf{x}. \tag{7}$$

The dual function $g(\mathbf{y}, \mathbf{s}) = \inf_{\mathbf{x}\geq 0} \mathcal{L}(\mathbf{x}, \mathbf{y}, \mathbf{s})$ is finite iff $\mathbf{c} - \mathbf{A}^\top\mathbf{y} - \mathbf{s} = \mathbf{0}$ with $\mathbf{s} \geq 0$, yielding $g(\mathbf{y}, \mathbf{s}) = \mathbf{b}^\top\mathbf{y}$. Thus the *dual* of (1) is

$$\max_{\mathbf{y}\in\mathbb{R}^m} \ \mathbf{b}^\top\mathbf{y} \quad \text{s.t.} \quad \mathbf{A}^\top\mathbf{y} \ \leq \ \mathbf{c}. \tag{8}$$

Weak duality. For any primal-feasible $\mathbf{x}$ and dual-feasible $\mathbf{y}$,

$$\mathbf{c}^\top\mathbf{x} \ \geq \ \mathbf{y}^\top\mathbf{A}\mathbf{x} \ = \ \mathbf{y}^\top\mathbf{b}. \tag{9}$$

Proof. $\mathbf{A}^\top\mathbf{y} \leq \mathbf{c}$ and $\mathbf{x} \geq 0$ imply $(\mathbf{c} - \mathbf{A}^\top\mathbf{y})^\top\mathbf{x} \geq 0$, giving (9). ■

KKT conditions (LP). Optimal $\left(\mathbf{x}^\star, \mathbf{y}^\star, \mathbf{s}^\star\right)$ satisfy

$$\text{Primal feasibility:} \quad \mathbf{A}\mathbf{x}^\star = \mathbf{b}, \quad \mathbf{x}^\star \geq 0. \tag{10}$$

$$\text{Dual feasibility:} \quad \mathbf{A}^\top\mathbf{y}^\star + \mathbf{s}^\star = \mathbf{c}, \quad \mathbf{s}^\star \geq 0. \tag{11}$$

$$\text{Complementary slackness:} \quad x_i^\star s_i^\star = 0, \quad \forall i = 1, \ldots, n. \tag{12}$$

These are sufficient and necessary for optimality because LPs are convex and satisfy Slater's condition when strictly feasible.

Complementary slackness (interpretation). For each i, either $x_i^\star > 0$ (then $s_i^\star = 0$ so the ith dual constraint is *tight*) or $s_i^\star > 0$ (then $x_i^\star = 0$ so the corresponding primal variable is at its lower bound).

Strong duality. If (1) is feasible and bounded, then (8) is feasible and $\min \mathbf{c}^\top\mathbf{x} = \max \mathbf{b}^\top\mathbf{y}$, with optimizers obeying KKT. Proof can be based on Farkas' lemma or separating hyperplanes (omitted for brevity).

Farkas' lemma (one form). Exactly one of the systems holds:

$$(\exists\,\mathbf{x} \geq 0 : \ \mathbf{A}\mathbf{x} = \mathbf{b}) \quad \text{or} \quad (\exists\,\mathbf{y} : \ \mathbf{A}^\top\mathbf{y} \geq 0, \ \mathbf{b}^\top\mathbf{y} < 0),$$

which underpins feasibility certificates and strong duality.

1.7 Simplex Optimality via Duality

At a basis $\mathcal{B}$, the multiplier $\mathbf{y}$ from (3) satisfies $\mathbf{A}_{\mathcal{B}}^\top \mathbf{y} = \mathbf{c}_{\mathcal{B}}$ and hence $\mathbf{s} = \mathbf{c} - \mathbf{A}^\top \mathbf{y}$ has $\mathbf{s}_{\mathcal{B}} = \mathbf{0}$ and $\mathbf{s}_{\mathcal{N}} = \bar{\mathbf{c}}_{\mathcal{N}}$ (reduced costs). Thus the simplex optimality test (5) is exactly *dual feasibility*; primal feasibility holds by construction and complementary slackness holds because $x_{\mathcal{N}} = 0$.

1.8 Interior-Point (Barrier) Derivation

Impose $x_i > 0$ via a logarithmic barrier and solve the barrier subproblem

$$\min_{\mathbf{x}>0} \; \mathbf{c}^\top \mathbf{x} \; - \; \mu \sum_{i=1}^{n} \ln x_i \quad \text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \qquad \mu > 0. \tag{13}$$

The Lagrangian of (13) with multiplier $\mathbf{y}$ for $\mathbf{A}\mathbf{x} = \mathbf{b}$ is

$$\mathcal{L}_\mu(\mathbf{x}, \mathbf{y}) = \mathbf{c}^\top \mathbf{x} - \mu \sum_i \ln x_i + \mathbf{y}^\top (\mathbf{b} - \mathbf{A}\mathbf{x}).$$

Stationarity yields

$$\nabla_{\mathbf{x}} \mathcal{L}_\mu = \mathbf{c} - \mathbf{A}^\top \mathbf{y} - \mu \, \mathbf{X}^{-1} \mathbf{1} = \mathbf{0},$$

i.e., with $\mathbf{s} \triangleq \mathbf{c} - \mathbf{A}^\top \mathbf{y}$,

$$\mathbf{X}\mathbf{S}\mathbf{1} = \mu \, \mathbf{1}, \qquad \mathbf{x} > 0, \; \mathbf{s} > 0. \tag{14}$$

Together with primal/dual feasibility ($\mathbf{A}\mathbf{x} = \mathbf{b}, \; \mathbf{A}^\top \mathbf{y} + \mathbf{s} = \mathbf{c}$), (14) defines the *central path.* As $\mu \downarrow 0$, $(\mathbf{x}, \mathbf{y}, \mathbf{s})$ converges to KKT of the LP.

Primal–dual Newton step. Linearize the KKT system

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{A}^\top \mathbf{y} + \mathbf{s} = \mathbf{c}, \quad \mathbf{X}\mathbf{S}\mathbf{1} = \mu \mathbf{1}$$

around $(\mathbf{x}, \mathbf{y}, \mathbf{s})$ to obtain for increments $(\Delta\mathbf{x}, \Delta\mathbf{y}, \Delta\mathbf{s})$:

$$\begin{bmatrix} \mathbf{0} & \mathbf{A} & \mathbf{0} \\ \mathbf{A}^\top & \mathbf{0} & \mathbf{I} \\ \mathbf{S} & \mathbf{0} & \mathbf{X} \end{bmatrix} \begin{bmatrix} \Delta\mathbf{x} \\ \Delta\mathbf{y} \\ \Delta\mathbf{s} \end{bmatrix} = \begin{bmatrix} \mathbf{b} - \mathbf{A}\mathbf{x} \\ \mathbf{c} - \mathbf{A}^\top \mathbf{y} - \mathbf{s} \\ \mu\mathbf{1} - \mathbf{X}\mathbf{S}\mathbf{1} \end{bmatrix}. \tag{15}$$

Eliminate $\Delta\mathbf{s} = -\Delta(\mathbf{A}^\top \mathbf{y}) + \Delta\mathbf{c}$ or, more commonly, from the third block: $\Delta\mathbf{s} = \mathbf{X}^{-1}(\mu\mathbf{1} - \mathbf{X}\mathbf{S}\mathbf{1} - \mathbf{S}\Delta\mathbf{x})$, and solve the *normal equations*

$$\mathbf{A}\mathbf{X}\mathbf{S}^{-1}\mathbf{A}^\top \, \Delta\mathbf{y} = \mathbf{A}\big(\mathbf{X}\mathbf{S}^{-1}(\mu\mathbf{1} - \mathbf{X}\mathbf{S}\mathbf{1}) + \mathbf{x} - \mathbf{X}\mathbf{S}^{-1}\mathbf{s}\big)$$

for $\Delta\mathbf{y}$, then recover $\Delta\mathbf{s}$ and $\Delta\mathbf{x}$. Choose step lengths to keep $(\mathbf{x}, \mathbf{s}) > 0$ and reduce μ geometrically.

1.9 Algorithms

Algorithm (Primal Simplex).

1. **Initialization:** Find an initial BFS (e.g., by Phase I on $\min \mathbf{1}^\top \mathbf{z}$ s.t. $\mathbf{A}\mathbf{x} + \mathbf{I}\mathbf{z} = \mathbf{b}, \; \mathbf{x}, \mathbf{z} \geq 0$).
2. **Iterate:** With basis $\mathcal{B}$ (matrix $\mathbf{B}$), compute $\mathbf{y}^\top = \mathbf{c}_{\mathcal{B}}^\top \mathbf{B}^{-1}$ and reduced costs $\bar{\mathbf{c}}_{\mathcal{N}} = \mathbf{c}_{\mathcal{N}} - \mathbf{N}^\top \mathbf{y}$.
3. If $\bar{\mathbf{c}}_{\mathcal{N}} \geq 0$, **stop** (optimal). Else pick entering q with $\bar{c}_q < 0$.
4. Compute direction $\mathbf{d}_{\mathcal{B}} = -\mathbf{B}^{-1}\mathbf{a}_q$ and step $\theta^\star$ via (6); pivot leaving p; update basis; repeat.

Algorithm (Primal–Dual Interior-Point).

1. **Initialization:** Find strictly feasible $(\mathbf{x}^{(0)}, \mathbf{y}^{(0)}, \mathbf{s}^{(0)})$ with $\mathbf{x}^{(0)}, \mathbf{s}^{(0)} > \mathbf{0}$; set $\mu = \frac{\mathbf{x}^\top \mathbf{s}}{n}$.
2. **Repeat:** Solve Newton system (15) for $(\Delta\mathbf{x}, \Delta\mathbf{y}, \Delta\mathbf{s})$ with a centering parameter; take step $(\mathbf{x}, \mathbf{y}, \mathbf{s}) \leftarrow (\mathbf{x}, \mathbf{y}, \mathbf{s}) + \alpha(\Delta\mathbf{x}, \Delta\mathbf{y}, \Delta\mathbf{s})$ with maximal $\alpha \in (0, 1]$ preserving positivity.
3. **Update:** Reduce $\mu \leftarrow \sigma \frac{\mathbf{x}^\top \mathbf{s}}{n}$ for $\sigma \in (0, 1)$ and repeat until primal/dual residuals and complementarity $\mathbf{x}^\top \mathbf{s}$ are below tolerance.

1.10 Summary of Variables and Their Dimensions

- $\mathbf{A} \in \mathbb{R}^{m\times n}$: constraint matrix; $\mathrm{rank}(\mathbf{A}) \le m$.
- $\mathbf{b} \in \mathbb{R}^m$: right-hand side (RHS); $\mathbf{c} \in \mathbb{R}^n$: cost vector.
- **Primal:** $\mathbf{x} \in \mathbb{R}^n$ (decision variables), with $\mathbf{x} \ge 0$.
- **Dual:** $\mathbf{y} \in \mathbb{R}^m$ (dual multipliers, free), $\mathbf{s} \in \mathbb{R}^n$ (dual slacks, $\mathbf{s} \ge 0$).
- **Basis objects:** $\mathcal{B} \subset \{1, \ldots, n\}$, $|\mathcal{B}| = m$; $\mathbf{B} = \mathbf{A}_{(:,\mathcal{B})} \in \mathbb{R}^{m\times m}$; $\mathcal{N}$ and $\mathbf{N}$ are complements.
- **BFS:** $\mathbf{x}_\mathcal{N} = \mathbf{0}$, $\mathbf{x}_\mathcal{B} = \mathbf{B}^{-1}\mathbf{b}$.
- **Simplex multipliers:** $\mathbf{y}^\top = \mathbf{c}_\mathcal{B}^\top \mathbf{B}^{-1}$; reduced costs $\bar{\mathbf{c}}_\mathcal{N} = \mathbf{c}_\mathcal{N} - \mathbf{N}^\top \mathbf{y}$.
- **KKT (LP):** $\mathbf{A}\mathbf{x} = \mathbf{b}$, $\mathbf{A}^\top \mathbf{y} + \mathbf{s} = \mathbf{c}$, $\mathbf{x} \ge 0$, $\mathbf{s} \ge 0$, $\mathbf{X}\mathbf{S}\mathbf{1} = \mathbf{0}$ at optimum.
- **Barrier/IPM:** central path $\mathbf{X}\mathbf{S}\mathbf{1} = \mu\mathbf{1}$ with $\mu > 0$; Newton system (15).

1.11 Summary

From first principles, an LP minimizes $\mathbf{c}^\top \mathbf{x}$ over a polyhedron $\{\mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \ge 0\}$. Extreme-point theory ensures that an optimum occurs at a basic feasible solution. The simplex method navigates among BFSs using reduced costs (dual feasibility) and a ratio test to preserve primal feasibility; optimality is certified by nonnegative reduced costs. The Lagrangian dual max $\mathbf{b}^\top \mathbf{y}$ subject to $\mathbf{A}^\top \mathbf{y} \le \mathbf{c}$ yields weak/strong duality and the KKT system with complementary slackness. Interior-point methods solve perturbed KKT equations along the central path $\mathbf{X}\mathbf{S}\mathbf{1} = \mu\mathbf{1}$, driving $\mu\downarrow 0$ to reach primal–dual optima. These perspectives (geometry, algebra, duality, and barriers) together provide complete, rigorous foundations and practical algorithms for linear programming.