Linear Programming Foundations: Simplex, Interior Point, and Duality Theory,

By Shaheer Ahmed Qureshi

Abstract

Abstract. Linear programming (LP) is the basis of optimization theory which has revolutionized post modern decision making in a variety of different fields. The following expository paper provides a rigorous explanation to three foundational pillars of linear programming the Simplex Method, Interior Point Methods, and Duality. This paper treats each topic through a unified five stage framework Intuit, Define, Mechanize, Mathematical Proof, and Bound. Ensuring both that there is complete conceptual clarity while simultaneously maintaining Mathematical rigor. The paper then provides a synthesis showing that Duality is not merely a third topic but the theoretical structure underlying both algorithms, the Simplex Method which implicitly constructs a dual solution at every iteration, while Interior Point which explicitly drive the primal-dual gap to zero along the central path. A concluding section acknowledges limitations and provides a directions for further study.

July 12, 2026

Contents

1 Introduction

Linear Programming (LP) is one of the most easily applicable and foundational tool in Optimization theory [6] . It has been able to bridge disciplines like mathematics, computer science, operations research, and engineering. At the center of it LP is about the optimization of a linear function like a profit or cost in accordance to a set of linear equality and inequality constraints. Contrary to it on surface simplicity, LP models are have the ability to represent a vast array of real-world decision-making problems, ranging from manufacturing schedules and logistics networks to machine learning pipelines and resource allocation in cloud computing environments.

Intuition

A practical scenario can serve as a perfect way to understand LP. Consider that a manufacturing plant produces two types of satellite components, component 1 (x_1) and component 2 (x_2), where each unit yields a profit of \$40 and \$30, respectively. Now The objective emerges is the desire to maximize total profit, which is written as the objective function $z = 40x_1 + 30x_2$. However, production cannot scale infinitely due to logistical physical constraints. The plant is restricted by a maximum of 60 available hours of technician assembly time and a strict material ceiling of 120 specialized units per week. If a single production of component 1 requires 2 hours of assembly and 4 specialized units, while a production of component 2 requires 3 hours of assembly and 2 specialized units, these resource boundaries translate directly into a system of linear inequalities: $2x_1 + 3x_2 \leq 60$ for labor capacity, and $4x_1 + 2x_2 \leq 120$ for unit supply, finally a logical domain is in order for our variables. Acknowledging that a factory cannot manufacture a negative quantity of hardware, which yields the non-negativity constraints $x_1, x_2 \geq 0$. Gathering these commands into a single, structure reveals the reader's first complete linear program:

$$\begin{aligned} \text{Maximize} \quad & z = 40x_1 + 30x_2 \\ \text{subject to} \quad & 2x_1 + 3x_2 \leq 60 \\ & 4x_1 + 2x_2 \leq 120 \\ & x_1, x_2 \geq 0 \end{aligned}$$

The nature of the feasible region ensures that if an optimal solution exists, it exist on a point from the ploytope produced as a result of the constraints This insight is central to the success of the simplex method and subsequent algorithmic innovations in LP. This example exemplifies the fact that real-world scenarios of limitations, the fundamental definition of a linear program becomes clear it is simply the language used to navigate choices under limitations.

Historical Context

In 1939, Soviet Union Mathematician Leonid Kantorovich published Matematicheskie metody organizatsi i planirovaniya proizvodstva ("Mathematical Methods for Organization and Planning of Production"). Which is accreditted as being the first study to identify the rigid mathematical structures underlying broad classes of scheduling problems, but the more widely acclaimed and uniform work on the subject being in US Air Force 1947 spearheaded by George Dantzig, where they produced the Simplex Method to quickly solve formerly impossible logistical and scheduling problems. Further more Dantzig later worked by John Von Neumenn to develop on the study of Duality in Linear Programming with Primal Value and Dual Value. [6]

Practical and Applied Perspective

Since its 20th century origin linear programming's methodology has evolved from exterior vertex tracking through Simplex method to advanced, interior-point polynomial-time algorithms capable of traversing the interior of complex polytopes [5]. Today LP serves as a foundational engine for decision-making and optimization across diverse disciplines. It supports logistics in operations research, produces market equilibrium models in economics, algorithms in computer science and real-time automation in systems engineering. Linear Programming in itself is very useful tool but it has set precedence for even more complex optimization systems like integer programming, nonlinear programming, and Semi definite programming. Now in the modern world which is data intensive and dependent on Artificial intelligence, Linear Programming plays a vital role, like when problems appear where there are hard set boundaries or there can be complexities. These can make it difficult to solve directly, so linear programming can come in simplify into linear models to get highly accurate approximations or even just establish a baseline for the best possible outcomes.

Scope of this Paper

The following paper aims to provide a rigorous explanation two methods that are essential in order to best apply Linear Programming, simplex Method and interior point Methods. Then also explain the theory of Duality and elaborate connection with the methods. It is not to emphasize the mathematical principles underlying each approach but also the algorithmic mechanics that make them practically useful. Finally the synthesis of these three pillars showing that Duality is the theoretical structure underlying why both the Simplex Method and Interior Point Methods converge to the same optimum.

2 The Simplex Method

2.1 Intuition

In order to intuit the simplex method we need to work around some of the basics of Linear Programming, consider the feasible region resulting from a Linear Program as a many-faced gemstone sitting in space a convex polytope. Which is bounded by flat faces, edges, and corners. In the preliminaries it has been established that if an optimum exists it lives at one of the corners of the said gemstone. This idea changes our mindset from continuously searching the infinitely many points in the feasible region to searching over finitely many corners.

So the intuitive picture is simple, observe a corner of the the feasible region look along each edge leading away from where you are. Now Some edges lead toward more ideal values, while others lead away from them. Walk along the edge that improves the objective the most then arrive at a new corner. Repeat the process because the gemstone has only can have finitely many corners and because at each step you only move to a better one until eventually you arrive somewhere with no improving edge left to walk along. So ultimately that corner is the optimum value.

This is the entire idea of the Simplex Method. What makes it nontrivial is not the idea of the method itself but making it precise, how do you represent a corner algebraically, how do you compute which edges are "improving," and how do you know when to stop. This is are factors of the method that are complex for the user, hence this precision is the subject of the next four stages.

2.2 Define

Definition 2.1 (Standard form). A linear program is in standard form if it is written as

$$\text{minimize} \quad c^T x$$

$$\text{subject to} \quad Ax = b, \quad x \geq 0$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, and $n > m$. We assume $\text{rank}(A) = m$, i.e., the rows of A are linearly independent.

Definition 2.2 (Basis). A basis is a choice of m linearly independent columns of A, indexed by a set $B \subset \{1, \ldots, n\}$ with $|B| = m$. The corresponding submatrix $A_B \in \mathbb{R}^{m \times m}$ is invertible.

Definition 2.3 (Basic solution). Given a basis B, the associated basic solution sets $x_i = 0$ for all $i \notin B$ (the nonbasic variables) and solves $A_B x_B = b$ for the basic variables $x_B = A_B^{-1} b$.

Definition 2.4 (Basic feasible solution, BFS). A basic solution is a basic feasible solution if it additionally satisfies $x_B \geq 0$.

Proposition 2.1. Basic feasible solutions correspond exactly to vertices of the feasible polytope $\{x : Ax = b, x \geq 0\}$.

Proof sketch: A vertex is a feasible point on the created polytope, which cannot be written as a strict convex combination of two other feasible points. It shows that a feasible point x is a vertex only the condition that columns of A corresponding to the nonzero entries of x are linearly independent, which is the condition defining a basic feasible solution (extended to allow some basic variables to be zero, the degenerate case). Hence this proposition acts as the bridge between the geometric intuition of 2.1 and the algebra followed by it confirming that "walking corner to corner" and "moving from one basic feasible solution to another" are the same operation viewed two ways.

Definition 2.5 (Reduced cost). For a nonbasic variable x_j, $j \notin B$, define the reduced cost

$$\bar{c}_j = c_j - c_B^T A_B^{-1} A_j$$

where A_j is the j-th column of A and c_B is the cost vector restricted to the basic indices. The reduced cost measures the rate of change of the objective if x_j is increased from zero while maintaining $Ax = b$.

2.3 Mechanize

Suppose at the current BFS some reduced cost $\bar{c}_j < 0$ for a nonbasic j, this means increasing x_j from zero would decrease the objective (we are minimizing). So the current solution is not optimal and we should bring x_j into the basis. As x_j increases from 0 the basic variables must adjust to maintain $Ax = b$. Write $x_B(t) = A_B^{-1}b - t \cdot A_B^{-1}A_j$ as $x_j = t$ increases. Each basic variable changes linearly in t. Feasibility ($x_B \geq 0$) requires t to stay below a threshold for every basic variable whose coefficient in this expression is positive — this is the minimum ratio test:

$$t^* = \min \left\{ \frac{(A_B^{-1}b)_i}{(A_B^{-1}A_j)_i} : (A_B^{-1}A_j)_i > 0 \right\}$$

The index i achieving this minimum identifies the basic variable that hits zero first and must leave the basis. This variable is called the leaving variable; x_j is the entering variable. The new basis $B' = B \setminus \{i\} \cup \{j\}$ defines the next BFS — geometrically, the next vertex along the edge we have just traversed.

The full algorithm:

1. Start with an initial BFS (basis B_0). Obtaining one may itself require a preliminary "Phase I" simplex run if no obvious starting basis is feasible — we note this without full elaboration, as it is a standard auxiliary procedure [8].
2. Compute reduced costs $\bar{c}_j$ for all nonbasic j.
3. If $\bar{c}_j \geq 0$ for all j, stop — the current BFS is optimal.
4. Otherwise, select an entering variable j with $\bar{c}_j < 0$ (commonly the most negative, Dantzig's rule).
5. Perform the minimum ratio test to find the leaving variable.
6. Update the basis, recompute A_B^{-1} (in practice via efficient updates, not full re-inversion), and return to step

Tie-breaking and degeneracy. If for instance the minimum ratio test results in a tie, the resulting BFS is degenerate (some basic variable equals zero) and naive pivoting can cycle indefinitely without making progress. Bland's rule (always choose the lowest-indexed eligible entering and leaving variable) provably prevents cycling at some cost to practical speed. We state this without proof the proof is a finite, somewhat tedious case analysis available in standard references [7, 8].

2.4 Mathematical Proofs

Theorem 2.1 (Simplex Optimality Criterion)**.** *A basic feasible solution with basis B is optimal if and only if $\bar{c}_j \geq 0$ for all nonbasic j.*

Proof. ($\Leftarrow$) Suppose $\bar{c}_j \geq 0$ for all nonbasic j. Let $\mathbf{x}$ be the current BFS and $\mathbf{y}$ any other feasible solution. Write $\mathbf{y} = \mathbf{x} + \mathbf{d}$ where $\mathbf{A}\mathbf{d} = \mathbf{0}$ and $x_i + d_i \geq 0$ for all i. Express $\mathbf{d}_B = -\mathbf{A}_B^{-1}\mathbf{A}_N\mathbf{d}_N$ where N denotes nonbasic indices. Then:

$$\mathbf{c}^\top\mathbf{y} - \mathbf{c}^\top\mathbf{x} = \mathbf{c}^\top\mathbf{d} = \sum_{j \notin B} \bar{c}_j\, d_j.$$

Since $x_j = 0$ for nonbasic j and feasibility forces $d_j \geq 0$, and since $\bar{c}_j \geq 0$ by hypothesis, every term is non-negative. Hence $\mathbf{c}^\top\mathbf{y} \geq \mathbf{c}^\top\mathbf{x}$ for every feasible $\mathbf{y}$, so $\mathbf{x}$ is optimal.

($\Rightarrow$) If some $\bar{c}_j < 0$, then increasing x_j along the pivot direction strictly decreases the objective while maintaining feasibility (for sufficiently small $t > 0$), contradicting optimality. □

Corollary 2.2. *If $\bar{c}_j > 0$ strictly for all nonbasic j, the optimal solution is unique.*

Corollary 2.3 (Finite termination, nondegenerate case)**.** *If every BFS is nondegenerate, the objective strictly decreases at every pivot. Since there are at most $\binom{n}{m}$ bases (a finite number) and no BFS is revisited, the algorithm terminates in finitely many steps.*

2.5 Complexity Bound

The Klee–Minty Result [1]

Klee and Minty (1972) [1] in 1972 showed that the Simplex Method under Dantzig's pivot rule, can require an *exponential* number of iterations in its worst case. They constructed for each dimension $d \geq 2$ a linear program whose feasible region is a combinatorially equivalent deformation of the d-dimensional unit cube $[0,1]^d$, on which the simplex path visits all $2^d - 1$ vertices before terminating. Specifically, they show:

$$H(d, 2d) \;\geq\; 2^d - 1,$$

where $H(d,n)$ denotes the maximum number of simplex pivots required over all nondegenerate LPs of class (d,n). This shed light on the fact that no polynomial bound exists for the worst-case iteration count of the Simplex Method under standard pivot rules.

In practice, however, simplex runs in far fewer steps on real and randomly generated instances, typically growing linearly or low-polynomially in m and n. A phenomenon explained by smoothed complexity analysis (Spielman and Teng, 2004). The Klee–Minty pathological family is what motivated the development of provably polynomial alternatives (interior point methods).

2.6 Worked Examples

Example 3.1 — Manufacturing Allocation (2 variables)

Problem. The factory from Section 1:

$$\max\ 6x_1 + 14x_2 \quad \text{s.t.} \quad \tfrac{1}{2}x_1 + 2x_2 \le 24,\ x_1 + 2x_2 \le 60,\ x_1, x_2 \ge 0.$$

Converting to standard form with slacks $x_3, x_4 \ge 0$:

$$\tfrac{1}{2}x_1 + 2x_2 + x_3 = 24, \quad x_1 + 2x_2 + x_4 = 60.$$

Iteration 0. Initial BFS: $B = \{3, 4\}$, $x_3 = 24$, $x_4 = 60$, $z = 0$. Reduced costs: $\bar{c}_1 = -6$, $\bar{c}_2 = -14$.
Iteration 1. Enter x_2 (most negative). Ratio test: $24/2 = 12$ (row 1), $60/2 = 30$ (row 2). Minimum is 12, so x_3 leaves. Pivot: new BFS $x_2 = 12$, $x_4 = 36$, $z = 168$.
Iteration 2. Updated reduced costs: $\bar{c}_1 = -2$. Enter x_1. Ratio test gives x_4 leaves. New BFS: $x_1 = 36$, $x_2 = 6$, $z = 294$.
Iteration 3. All reduced costs ≥ 0. **Optimal:** $x_1 = 36$, $x_2 = 6$, $z = 294$.

Example 3.2 — Diet Problem (3 variables)

Problem. Minimise cost subject to nutritional lower bounds:

$$\min\ 2x_1 + 3x_2 + x_3 \quad \text{s.t.} \quad x_1 + x_2 \ge 4,\ x_1 + x_3 \ge 3,\ x_1, x_2, x_3 \ge 0.$$

Introduce surplus variables $s_1, s_2 \ge 0$ and artificial variables $a_1, a_2 \ge 0$ (Phase I). After Phase I establishes feasibility, Phase II runs the simplex method on the original cost function.
Phase II result (details omitted for brevity): $x_1 = 0$, $x_2 = 4$, $x_3 = 3$, $z = 15$.
Reduced costs at termination: $\bar{c}_1 = 2 - (1 \cdot y_1 + 1 \cdot y_2) = 2 - 2 = 0$ confirming optimality via Theorem 2.1. Shadow prices $y_1 = 3$, $y_2 = 1$ satisfy dual feasibility (see Section 5).

2.7 Real-World Scenarios

Scenario 3.1 — Airline Crew Scheduling

Airlines management are supposed to assign crews members to flights subject to FAA rest requirements, aircraft type ratings, base constraints and union rules (constraints). A major carrier then constructs this as an LP with hundreds of thousands of variables (one per possible crew pairing) and tens of thousands of constraints (one per flight leg), here arrives the use of the Simplex Method. Exploiting the extremely sparse structure of the constraint matrix (each pairing covers only a few legs), solves these problems in minutes in operational use. The warm-start property of simplex is critical here and in the scenario when a flight is delayed and the problem is re-solved with a slightly perturbed right-hand side, the previous optimal basis provides a starting point that converges in very few additional pivots.

Scenario 3.2 — Portfolio Optimisation (Mean–Variance)

A fund manager selects weights $x_j \geq 0$ for n assets to maximize expected return $\mathbf{c}^\top \mathbf{x}$ subject to a budget constraint $\sum_j x_j = 1$ and upper bounds on individual positions. When the covariance matrix is replaced by a linear approximation (for example a factor model fixing variance) the problem transforms into a LP, solvable by the Simplex Method. The simplex tableau's reduced costs directly read out as each asset's marginal contribution to the objective. Providing the portfolio manager with immediate economic interpret-ability.

3 Interior Point Methods

3.1 Intuit

In contrast to the Simplex Method, which moves along a vertexes and edges, the interior point methods abandon the boundary altogether. Now visualize a scenario starting at a point deep inside the gemstone, nowhere near any surface and then moving smoothly through its interior, Carefully tracing a curved path that bends ever closer to the optimal vertex without ever touching the boundary until the very end, but the question arise why do this? it is simply because the boundary is where the combinatorial difficulty lives. The Klee–Minty construction is pathological precisely because it forces the boundary-walker to visit exponentially many corners. The interior of the polytope is smooth, no corners and no edges. Just an open convex set. Calculus, gradients, Newton's method, smooth optimization is far better suited to a smooth interior than to a faceted boundary. Interior point methods exploit this by converting the combinatorial problem of which vertex into the analytic problem of follow this curve.

3.2 Define

Definition 3.1 (Logarithmic barrier function). For the standard-form LP's nonnegativity constraints, define the barrier function

$$B(x) = -\sum_{j=1}^{n} \log(x_j)$$

defined on the strict interior of the feasible region ($x_j > 0$ for all j). Note $B(x) \to +\infty$ as any $x_j \to 0^+$, so B penalizes proximity to the boundary increasingly severely.

Definition 3.2 (Barrier problem). For a parameter $\mu > 0$, define the barrier-modified problem

$$\begin{aligned} &\text{minimize} \quad c^T x + \mu B(x) \\ &\text{subject to} \quad Ax = b \end{aligned}$$

This is now a smooth, equality-constrained problem — the inequality constraints $x \geq 0$ have been absorbed into the objective via the barrier term, leaving only the linear equality constraint, which can be handled by standard Lagrangian methods.

Definition 3.3 (Central path). Let $x^*(\mu)$ denote the unique minimizer of the barrier problem for a given $\mu > 0$ (existence and uniqueness follow from strict convexity of B on the interior — a fact we state without full proof, though it follows from B's Hessian being positive definite). The set $\{x^*(\mu) : \mu > 0\}$ is called the central path [9].

3.3 Mechanize

Deriving the optimality conditions. Introduce a Lagrange multiplier $y \in \mathbb{R}^m$ for the equality constraint $Ax = b$. The Lagrangian is

$$\mathcal{L}(x, y) = c^T x + \mu B(x) - y^T (Ax - b)$$

Setting $\partial \mathcal{L}/\partial x_j = 0$ for each j gives

$$c_j - \frac{\mu}{x_j} - (A^T y)_j = 0, \quad \text{i.e.,} \quad c_j - (A^T y)_j = \frac{\mu}{x_j}$$

Define $s = c - A^T y$ (this will turn out to be exactly the dual slack vector, anticipating Section 4). The condition becomes $s_j x_j = \mu$ for every j — written compactly as

$$XSe = \mu e$$

where $X = \text{diag}(x_1, \ldots, x_n)$, $S = \text{diag}(s_1, \ldots, s_n)$, and e is the all-ones vector. Together with the original equality constraint $Ax = b$ and the dual feasibility condition $A^T y + s = c$, this gives the full system of central path equations:

$$Ax = b, \quad A^T y + s = c, \quad XSe = \mu e, \quad x, s > 0$$

Newton's method on the central path. Given a current iterate (x, y, s) and a target μ, we solve this nonlinear system approximately by linearizing. Newton's method replaces the nonlinear equation $XSe = \mu e$ with its first-order (linear) approximation around the current point, producing a linear system in the search directions $(\Delta x, \Delta y, \Delta s)$:

$$\begin{aligned} A\Delta x &= b - Ax \\ A^T \Delta y + \Delta s &= c - A^T y - s \\ S\Delta x + X\Delta s &= \mu e - XSe \end{aligned}$$

This is a linear system solvable by standard linear algebra (typically reduced to a smaller "normal equations" system in Δy for computational efficiency, which we mention without full derivation [9]). A step length $\alpha \in (0, 1]$ is chosen to ensure $(x + \alpha\Delta x, s + \alpha\Delta s)$ remains strictly positive, and μ is decreased geometrically (e.g., multiplied by some $\sigma \in (0, 1)$) at each outer iteration.

The full algorithm:

1. Initialize an interior point (x^0, y^0, s^0) with $x^0, s^0 > 0$ (obtaining one is a nontrivial auxiliary step, analogous to Phase I in simplex).
2. Choose initial μ (often related to $(x^0)^T s^0/n$) and a reduction factor σ.
3. Solve the Newton system above for the search direction.
4. Choose a step length preserving strict positivity; update (x, y, s).
5. Reduce $\mu \leftarrow \sigma\mu$.
6. If the duality gap $x^T s$ is below a tolerance ε, stop. Otherwise return to step

3.4 Mathematical Proof

Proposition 3.1 (Existence and Uniqueness on Central Path)**.** *For each $\mu > 0$, the barrier problem has a unique minimiser $\mathbf{x}^*(\mu)$ in the strict interior of P, provided P has nonempty interior and is bounded.*

Proof. The Hessian of $B(\mathbf{x})$ is $\operatorname{diag}(x_1^{-2}, \ldots, x_n^{-2})$, which is positive definite for all $\mathbf{x} > \mathbf{0}$. Hence $\mathbf{c}^\top\mathbf{x} + \mu B(\mathbf{x})$ is strictly convex on the interior. A strictly convex function on a convex domain has at most one minimiser. Boundedness of P and the fact that $B(\mathbf{x}) \to +\infty$ as any component approaches zero guarantee the minimiser exists and lies in the strict interior. □

Theorem 3.2 (Polynomial-Time Convergence)**.** *Under a short-step or predictor-corrector strategy for μ, the interior point method (Algorithm **??**) produces an iterate with duality gap below ε in*

$$O\left(\sqrt{n} \cdot \ln \frac{1}{\varepsilon}\right)$$

iterations, each requiring $O(n^3)$ arithmetic operations. This gives an overall polynomial-time algorithm in the problem dimension.

Proof sketch. The key insight, due to Karmarkar [2], is that the *potential function*

$$f(\mathbf{x}) = \sum_j \ln\left(\frac{\mathbf{c}^\top\mathbf{x}}{x_j}\right)$$

decreases by a constant $\delta > 0$ at every step of the algorithm. Since f starts at $O(n \ln n)$ and the algorithm terminates when $f \leq -O(q)$ (corresponding to duality gap $\leq 2^{-q}$), the number of steps is $O(n \ln(1/\varepsilon))$. Karmarkar's projective transformation shows this decrease can always be achieved, giving the $\sqrt{n}$ factor in the iteration bound for short-step methods. □

Interior Point vs. Simplex: Complexity Comparison

Property	Simplex	Interior Point
Worst-case iterations	Exponential [1]	$O(\sqrt{n}\log(1/\varepsilon))$
Typical iterations (practice)	$O(m)$ to $O(3m)$	$O(30)$–$O(100)$
Iteration cost	$O(m^2n)$ (sparse)	$O(n^3)$ (dense)
Warm starting	Natural	Difficult
Large-scale sparse	Excellent	Moderate
Large-scale dense	Moderate	Excellent
Trajectory	Boundary (vertices)	Interior

The polynomial guarantee of interior point methods resolved the theoretical gap exposed by Klee and Minty, but neither method dominates the other in practice. Their strengths are complementary.

3.5 Worked Examples

Example 4.1 — Central Path Tracking (2 variables)

Problem.

$$\min\ -x_1 - x_2 \quad \text{s.t.} \quad x_1 + x_2 \le 1,\ x_1, x_2 \ge 0.$$

Standard form: $x_1 + x_2 + s = 1$, $\mathbf{x} = (x_1, x_2, s)^\top \ge \mathbf{0}$.
Central path. Setting $s_j x_j = \mu$ for each variable and solving with the equality constraint:

$$x_1^*(\mu) = x_2^*(\mu) = \frac{1}{2} - \mu + O(\mu^2), \quad s^*(\mu) = 2\mu + O(\mu^2) \text{ (as } \mu \to 0).$$

As $\mu \to 0$, the path converges to $(1/2, 1/2, 0)$, the optimal vertex. The interior point algorithm traces this smooth curve from the initial interior point to the optimum, never touching the boundary.
Newton step from $(0.4, 0.4, 0.2)$**.** Here $\mu = 0.4 \times 0.4 + 0.4 \times 0.4 + 0.2 \times 0.2 = 0.36$ (unnormalised). After one Newton iteration toward $\sigma\mu$, the iterate moves interior-wise toward $(0.5, 0.5, 0)$, demonstrating the algorithm's interior trajectory.

Example 4.2 — Minimum Cost Flow

Consider a transportation problem: supply nodes S_1, S_2, demand nodes D_1, D_2, and unit shipping costs. The LP has 4 variables (flows on each arc) and 4 constraints (supply/demand balance). The barrier problem at $\mu = 0.1$ yields central path conditions:

$$c_{ij} - y_i^{\text{supply}} + y_j^{\text{demand}} = \frac{\mu}{x_{ij}},$$

which can be interpreted as: at the central path, the reduced cost of each arc equals μ/x_{ij} (positive and decreasing to zero as $\mu \to 0$). Interior point methods are particularly effective for large transportation problems where the network structure makes the Newton system highly structured and efficiently solvable.

3.6 Real-World Scenarios

Scenario 4.1 — Power Grid Optimal Dispatch

Independent System Operators are solving Security-Constrained Economic Dispatch linear programs every five minutes to minimize generation cost across thousands of generators. Subject to transmission limits, ramp constraints and security contingency requirements. These linear programs often have millions of variables and hundreds of thousands of constraints. Interior point methods are preferred here because their polynomial worst-case guarantee provides predictable computation times under tight real-time deadline and unlike simplex they do not degrade on dense constraint matrices arising from contingency analysis, also the central path provides a smooth sequence of solutions that can be leveraged for sensitivity analysis and pricing of transmission congestion.

Scenario 4.2 — Machine Learning: Support Vector Machines

Training a Support Vector Machine with a linear kernel requires solving a quadratic program. When the kernel is linearized or the problem is cast in its dual form it reduces to an LP. Interior point methods become the standard approach to it because the SVM dual LP is dense, where interior point's $O(n^3)$ per-iteration cost is competitive with simplex's much-higher effective cost on dense problems. Interior point naturally produces soft-margin classifiers by keeping all slack variables strictly positive and the duality gap provides a natural stopping criterion aligned with the statistical generalization guarantee of the classifier.

4 Duality

4.1 Motivation: Bounding from Below

Every primal feasible point gives an upper bound on the minimum objective. Is there a way to construct lower bounds *without* solving the problem? Duality answers this question [4]. We seek a vector $\mathbf{y} \in \mathbb{R}^m$ such that $\mathbf{y}^\top \mathbf{b}$ is a guaranteed lower bound on $\mathbf{c}^\top \mathbf{x}$ for every primal-feasible $\mathbf{x}$.

4.2 Deriving the Dual

Since $\mathbf{A}\mathbf{x} = \mathbf{b}$ for any feasible $\mathbf{x}$, we have $\mathbf{y}^\top \mathbf{b} = \mathbf{y}^\top \mathbf{A}\mathbf{x} = (\mathbf{A}^\top \mathbf{y})^\top \mathbf{x}$. For this to be a lower bound on $\mathbf{c}^\top \mathbf{x}$ for all $\mathbf{x} \geq \mathbf{0}$, we need $(\mathbf{c} - \mathbf{A}^\top \mathbf{y})^\top \mathbf{x} \geq 0$ for all $\mathbf{x} \geq \mathbf{0}$, which holds if and only if $\mathbf{A}^\top \mathbf{y} \leq \mathbf{c}$. Among all such $\mathbf{y}$, we want the *tightest* (largest) lower bound. This gives the *dual problem*:

$$\max_{\mathbf{y}\in\mathbb{R}^m} \mathbf{y}^\top\mathbf{b} \quad \text{subject to} \quad \mathbf{A}^\top\mathbf{y} \le \mathbf{c}.$$

The original problem is the *primal*; this is the *dual*. Note that the dual of the dual is the primal—a fundamental symmetry [7, 8].

4.3 Shadow Prices and Economic Interpretation

The dual variables $\mathbf{y}$ have a concrete economic meaning as *shadow prices* or *marginal values* of the resources $\mathbf{b}$ [3]. In the factory example, $y_1 = 11$ and $y_2 = \frac{1}{2}$: one additional metalworking-day is worth \$1,100, one additional woodworking-day is worth \$50. These are the breakeven rents at which the factory would exactly profit from leasing additional capacity. The dual problem asks: what are the lowest resource prices that rationalise all production decisions? This is precisely the linear program

$$\min_{y_1,y_2\ge 0} 24y_1 + 60y_2 \quad \text{s.t.} \quad \tfrac{1}{2}y_1 + y_2 \ge 6,\ 2y_1 + 2y_2 \ge 14,\ y_1 + 4y_2 \ge 13,$$

whose solution $y_1 = 11$, $y_2 = \frac{1}{2}$, $v = 294$ coincides exactly with the primal optimum.

4.4 The Three Duality Theorems

Theorem 4.1 (Weak Duality)**.** *For any primal-feasible* $\mathbf{x}$ *and dual-feasible* $\mathbf{y}$*:*

$$\mathbf{y}^\top\mathbf{b} \;\le\; \mathbf{c}^\top\mathbf{x}.$$

Proof. Since $\mathbf{A}\mathbf{x} = \mathbf{b}$, $\mathbf{x} \ge \mathbf{0}$, and $\mathbf{A}^\top\mathbf{y} \le \mathbf{c}$:

$$\mathbf{y}^\top\mathbf{b} = \mathbf{y}^\top\mathbf{A}\mathbf{x} = (\mathbf{A}^\top\mathbf{y})^\top\mathbf{x} \le \mathbf{c}^\top\mathbf{x},$$

where the last inequality holds because $\mathbf{x} \ge \mathbf{0}$ and $(\mathbf{c} - \mathbf{A}^\top\mathbf{y}) \ge \mathbf{0}$ (dual feasibility), so each term $(c_j - (\mathbf{A}^\top\mathbf{y})_j)x_j \ge 0$. □

Theorem 4.2 (Strong Duality)**.** *If the primal has an optimal solution* $\mathbf{x}^*$*, then the dual has an optimal solution* $\mathbf{y}^*$ *and* $\mathbf{c}^\top\mathbf{x}^* = (\mathbf{y}^*)^\top\mathbf{b}$*.*

Proof. At the optimal simplex basis B (from Theorem 2.1), define $\mathbf{y}^* = (\mathbf{c}_B^\top\mathbf{A}_B^{-1})^\top$. We verify:

(i) *Dual feasibility*: For each j, the j-th dual constraint is $(\mathbf{A}^\top\mathbf{y}^*)_j \le c_j$, i.e., $c_j - \mathbf{c}_B^\top\mathbf{A}_B^{-1}\mathbf{A}_j \ge 0$, which is exactly $\bar{c}_j \ge 0$. This holds by Theorem 2.1.

(ii) *Equality of objectives*: $(\mathbf{y}^*)^\top\mathbf{b} = \mathbf{c}_B^\top\mathbf{A}_B^{-1}\mathbf{b} = \mathbf{c}_B^\top\mathbf{x}_B^* = \mathbf{c}^\top\mathbf{x}^*$, since nonbasic components of $\mathbf{x}^*$ are zero.

By Theorem 4.1, $\mathbf{y}^*$ is dual-optimal. □

Theorem 4.3 (Complementary Slackness)**.** $\mathbf{x}^*$ *and* $\mathbf{y}^*$ *are simultaneously optimal for the primal and dual if and only if for every* $j = 1, \ldots, n$*:*

$$x_j^* \cdot (c_j - (\mathbf{A}^\top\mathbf{y}^*)_j) = 0.$$

Equivalently: for each j, either $x_j^* = 0$ (variable is zero) or $c_j = (\mathbf{A}^\top \mathbf{y}^*)_j$ (dual constraint is tight).

Proof. From the proof of Theorem 4.1, the duality gap is:

$$\mathbf{c}^\top \mathbf{x} - \mathbf{y}^\top \mathbf{b} = \sum_{j=1}^{n} (c_j - (\mathbf{A}^\top \mathbf{y})_j)\, x_j.$$

Each term $(c_j - (\mathbf{A}^\top \mathbf{y})_j)x_j \geq 0$ by feasibility. The gap is zero (i.e., strong duality holds) if and only if every term is zero, which gives the complementary slackness condition. The converse (CS conditions imply optimality) follows because CS implies zero duality gap, which by Theorem 4.1 implies simultaneous optimality. □

4.5 Duality and the Simplex Method

The dual variables $\mathbf{y}^* = \mathbf{c}_B^\top \mathbf{A}_B^{-1}$ are generated for free from the optimal simplex basis. In the factory example, the shadow prices $y_1 = 11$ and $y_2 = \frac{1}{2}$ emerge directly from the final simplex tableau as the negatives of the reduced costs of the slack variables. The simplex stopping condition (all reduced costs ≥ 0) *is* dual feasibility. Thus simplex simultaneously solves both the primal and dual problem, producing a complementary pair at termination.

4.6 Duality and the Interior Point Method

The central path equations include $\mathbf{A}^\top \mathbf{y} + \mathbf{s} = \mathbf{c}$, which is precisely dual feasibility. The duality gap at any interior point iterate is:

$$\mathbf{c}^\top \mathbf{x} - \mathbf{y}^\top \mathbf{b} = \mathbf{x}^\top \mathbf{s} = n\mu.$$

As $\mu \to 0$, the duality gap $\to 0$: the algorithm explicitly drives the gap to zero along the central path. Interior point methods are therefore *primal-dual* in that they maintain feasibility for both the primal and dual simultaneously and measure progress via the duality gap [2].

5 Synthesis

Now we can provide the paper's central claim precisely having provided all the aforementioned information in definitions, proofs, and algorithms.

Central Claim

The Simplex Method and Interior Point Methods are not independent algorithms that happen to solve the same problem but they are two different routes to the same destination. Which is th exact satisfaction of the complementary slackness conditions and Duality is the map that shows why that destination is unique and how each route finds it.

5.1 Connection 1: Simplex Constructs a Dual Certificate Implicitly

The Simplex Method never explicitly solves the dual LP. Yet at termination it produces a dual-optimal solution as a byproduct of solving the primal. At the optimal basis B, the vector

$$\mathbf{y}^* = \left(\mathbf{c}_B^\top \mathbf{A}_B^{-1}\right)^\top$$

is computed using quantities already available from the final tableau. Theorem 4.2 shows this $\mathbf{y}^*$ is dual feasible and achieves $(\mathbf{y}^*)^\top \mathbf{b} = \mathbf{c}^\top \mathbf{x}^*$.

More strikingly: the simplex stopping condition is not an arbitrary halting rule. By Definition **??**:

$$\bar{c}_j = c_j - (\mathbf{A}^\top \mathbf{y}^*)_j.$$

The condition $\bar{c}_j \geq 0$ for all nonbasic j is precisely the dual feasibility condition $\mathbf{A}^\top \mathbf{y}^* \leq \mathbf{c}$ (the constraint is satisfied for all j, including the basic ones where $\bar{c}_j = 0$ by construction). The Simplex Method terminates when it has found a dual certificate—it just reads it off the tableau without the user needing to know the dual problem exists.

5.2 Connection 2: Interior Point Makes the Primal–Dual Relationship Explicit

Interior point methods make what simplex does implicitly into the explicit architecture of the algorithm. At every iterate, the method maintains:

- A primal-feasible $\mathbf{x} > 0$ (satisfying $\mathbf{A}\mathbf{x} = \mathbf{b}$, $\mathbf{x} > \mathbf{0}$).
- A dual-feasible $(\mathbf{y}, \mathbf{s})$ (satisfying $\mathbf{A}^\top \mathbf{y} + \mathbf{s} = \mathbf{c}$, $\mathbf{s} > \mathbf{0}$).
- A duality gap $\mathbf{x}^\top \mathbf{s} = n\mu > 0$ that decreases at every step.

The algorithm terminates not because some combinatorial condition on a basis is met, but because the duality gap falls below a threshold. The central path is parameterised by $\mu = \mathbf{x}^\top \mathbf{s}/n$—a direct measure of the duality gap. Following the central path to $\mu = 0$ is, in geometric terms, following the duality gap to zero.

5.3 Connection 3: Both Algorithms Are Correct for the Same Reason

Both algorithms are guaranteed to find the global optimum, and neither success is a numerical accident. They terminate at points satisfying complementary slackness (Theorem 4.3), and the proof of strong duality (Theorem 4.2) guarantees that any such primal–dual pair is simultaneously optimal for both problems.

The two algorithms differ entirely in *how* they search:

Simplex	**Interior Point**
Boundary: vertex to vertex	Interior: smooth curve
Combinatorial pivots	Continuous Newton steps
Dual solution appears only at termination	Dual solution maintained at every iterate
Stopping: all $\bar{c}_j \geq 0$	Stopping: duality gap $\mathbf{x}^\top\mathbf{s} < \varepsilon$
Exponential worst case	Polynomial worst case

Both are searching for the unique certificate of optimality whose existence Theorem 4.2 guarantees in the abstract.

6 Conclusion

This paper has developed linear programming from a concrete factory problem through three interconnected pillars.

The **Simplex Method** transforms optimisation over an infinite feasible region into a finite search over vertices, justified by the Fundamental Theorem of LP. Its pivot operations walk the boundary of the feasible polytope, guided by reduced costs, terminating when the optimality criterion (Theorem 2.1) is met. Despite exponential worst-case behaviour [1], it remains dominant in practice for sparse structured problems.

Interior Point Methods overcome the theoretical deficiency by moving through the interior of the feasible region along the central path, a smooth trajectory parameterised by the duality gap. Karmarkar's projective transformation [2] guarantees a constant reduction in potential at every step, yielding the first polynomial-time algorithm with practical computational performance.

Duality is the theoretical structure that explains both. Weak duality provides a bound; strong duality proves it is tight; and complementary slackness characterises optimality as the mutual satisfaction of conditions derivable from either the primal or the dual. The simplex method satisfies these conditions combinatorially at the final vertex; interior point methods satisfy them analytically as $\mu \to 0$.

Limitations. This paper assumed primal feasibility and boundedness throughout. Degenerate cases (multiple optima, cycling in simplex, zero duality gap with no optimal solution) require additional theory. The barrier method as presented assumes a feasible interior starting point, which itself requires an auxiliary Phase I procedure. Nonlinear objectives break the linearity on which all three pillars rest.

Future directions. The natural next step is *Lyapunov stability theory* for nonlinear systems, where eigenvalue analysis generalises to energy-like functions that certify convergence without solving differential equations directly. In LP, the analogous extension is *semidefinite programming* (SDP), where the matrix variable replaces the vector $\mathbf{x}$ and duality theory extends with equal elegance. Integer programming, where variables are constrained to integers, breaks the clean LP duality theory and requires fundamentally new tools (branch-and-bound, cutting planes, Gomory cuts)—a natural and challenging next chapter.

Acknowledgment The Author would like to appreciate Trenton von Sosen and Dr Simon Rubenstein-Salzado for there assistance and guidance during the drafting of this paper

References

[1] V. Klee and G. J. Minty, *How Good Is the Simplex Algorithm?*, in O. Shisha (ed.), *Inequalities III*, Academic Press, New York, 1972, pp. 159–179.

[2] N. Karmarkar, *A New Polynomial-Time Algorithm for Linear Programming*, *Combinatorica*, **4**(4):373–395, 1984. Also published in *Proceedings of the 16th Annual ACM Symposium on Theory of Computing*, 1984, pp. 302–311.

[3] Applied Mathematics Programme, *Duality in Linear Programming* (Chapter 4), MIT Open Courseware, 2004.

[4] *Duality Theory in Linear Programming*, Course Notes (dual_121), 2024. Covering: the dual LP, the dual theorem, the complete slackness theorem, simplex tableau derivations, and the Farkas theorem.

[5] M. A. Schulze, *Linear Programming for Optimization*, Perceptive Scientific Instruments, Inc., 1998.

[6] S. J. Wright, *Optimization*, *Encyclopedia Britannica*, 21 May 2026. Available at: `https://www.britannica.com/science/optimization`. Accessed 3 July 2026.

[7] A. Schrijver, *Theory of Linear and Integer Programming*, Wiley-Interscience, Chichester, 1986.

[8] D. Bertsimas and J. N. Tsitsiklis, *Introduction to Linear Optimization*, Athena Scientific, Belmont, MA, 1997.

[9] S. J. Wright, *Primal-Dual Interior-Point Methods*, SIAM, Philadelphia, 1997.