Proceedings of CONF-CIAP 2026 Symposium: Applied Mathematics and Statistics
DOI: 10.54254/2753-8818/2026.CH31695

Algebraic Geometry in Optimization Theory

Enli Chen

Overseas Chinese Academy of Concord Suzhou, Suzhou, China
badmathgamer@outlook.com

Abstract. This essay traces the historical development of optimization, focusing on one-variable problems and their applications. We discuss the role of non-negative polynomials in optimization and carefully examine a classical result by Motzkin. Building on this result, we construct a new polynomial that is non-negative but not a sum-of-squares polynomial, propose a method to check whether a polynomial is sum-of-squares, and explore the relationship between the set of sum-of-squares polynomials and the set of non-negative polynomials.

Keywords: Optimization, Sum of squares, Non-negative polynomial, Gram matrix, Convex geometry

1. Introduction

1.1. Importance of optimization

Optimization stands as a cornerstone of progress in modern systems, spanning disciplines from engineering to economics. At its core, optimization seeks to maximize efficiency, minimize waste, and allocate resources intelligently. Traditional approaches, such as linear programming and gradient-based algorithms, have achieved remarkable success in structured, convex settings. However, the growing complexity of real-world problems—characterized by high dimensionality, non-convexity, and nonlinear constraints—demands innovative methodologies to overcome computational bottlenecks and ensure global optimality. This urgency motivates the exploration of unconventional mathematical frameworks, particularly cutting-edge ones, to redefine the boundaries of optimization theory.

1.2. What is algebraic geometry

Geometry, traditionally defined as the study of spatial properties and relational structures, is a multifaceted discipline that extends far beyond the analysis of shapes. While its classical focus lies in characterizing forms—from Euclidean polygons to differential manifolds—the methodologies underpinning geometric inquiry are strikingly diverse. A perennial question in mathematics concerns equivalence: under what conditions can distinct entities be deemed fundamentally identical? This question permeates all branches of geometry, though its interpretation varies by context. Algebraic geometry, a branch of mathematics studying solutions to polynomial equations through geometric principles, offers profound insights into the structural properties of complex systems. By analyzing

DOI: 10.54254/2753-8818/2026.CH31695

algebraic varieties (geometric manifestations of equation systems) and their intrinsic symmetries, this field provides tools to characterize solution spaces with rigorous precision.

1.3. Great potential in the combination of algebraic geometry and optimization

The fusion of algebraic geometry and optimization holds transformative potential. Techniques such as Gröbner basis theory enable the systematic elimination of variables in polynomial systems, facilitating exact solutions to non-convex optimization problems. By leveraging the duality between algebraic constraints and geometric intuition, this interdisciplinary approach promises to unlock efficient, theoretically grounded solutions for 21st-century challenges, from quantum computing to climate modeling.

1.4. Main contribution

In this paper, we analyze how algebraic tools can be applied to solving optimization problems. Our main contribution is to clarify the relation between non-negative polynomials and an important subset of them: those that can be written as sums of squares (SOS) of real polynomials. More specifically, we provide a way to construct families of non-negative polynomials that are not SOS. We also develop a method that can be used to numerically compute the boundary of non-SOS polynomials within the cone of all non-negative polynomials.

2. Fundamental notions in algebra

Algebraic geometry and many optimization problems build on algebraic notions. In this section, we introduce a group of concepts that will be used extensively in the following parts.

Definition 2.1. A vector space over the real numbers $\mathbb{R}$ is a set V with two operations, scalar multiplication $\cdot$ and addition $+$, satisfying the following properties for all $u, v, w \in V$ and $c, d \in \mathbb{R}$:

1. $u + v \in V$.
2. $u + v = v + u$.
3. $(u + v) + w = u + (v + w)$.
4. There exists a special vector $0_V \in V$ such that $u + 0_V = u$ for all $u \in V$.
5. For every $u \in V$, there exists $w \in V$ such that $u + w = 0_V$.
6. $c \cdot v \in V$.
7. $(c + d) \cdot v = c \cdot v + d \cdot v$.
8. $c \cdot (u + v) = c \cdot u + c \cdot v$.
9. $(cd) \cdot v = c \cdot (d \cdot v)$.
10. $1 \cdot v = v$ for all $v \in V$.

Remark 2.2. Similarly, we can also define vector spaces over the complex numbers or over a general field.

The concept of a vector space is the central object studied in linear algebra. Many objects in the real world can be modeled as vector spaces over certain fields (not only over $\mathbb{R}$ or $\mathbb{C}$). Between vector spaces, we are particularly interested in linear maps.

Definition 2.3. Given vector spaces V_1 and V_2 over $\mathbb{R}$ (or $\mathbb{C}$), a linear map $f: V_1 \to V_2$ is a function satisfying the following properties:

1. $f(x_1 + x_2) = f(x_1) + f(x_2)$ for all $x_1, x_2 \in V_1$.
2. $f(\alpha x) = \alpha f(x)$ for all $x \in V_1$ and $\alpha \in \mathbb{R}$ (or $\mathbb{C}$).

DOI: 10.54254/2753-8818/2026.CH31695

3. $f(0_{V_1}) = 0_{V_2}$.

We refer to [1] for more notions in linear algebra such as the characteristic polynomial, determinant, and related topics.

Definition 2.4. Let V be a vector space over $\mathbb{R}$. A subset $C \subset V$, with $C \neq \emptyset$, is called a cone if $\alpha c \in C$ for all $c \in C$ and all $\alpha \geq 0$.

Cones are a central concept in this paper, since many optimization problems can be associated with certain cones.

Definition 2.5. For a field k , any expression of the form

$$f(x) = \sum_{i=0}^{n} a_i x^i = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n, \tag{1}$$

where $a_i \in k$ and $a_n \neq 0$, is called a polynomial over k with indeterminate x . The elements $a_0, a_1, \ldots, a_n$ are called the coefficients of f .

Definition 2.6. A topology on a set X is a collection τ of subsets of X , called open sets, satisfying the following axioms:

1. The empty set and X itself belong to τ .
2. Any arbitrary (finite or infinite) union of members of τ belongs to τ .
3. The intersection of any finite number of members of τ belongs to τ .

This definition of a topology is the most commonly used one. The set τ of open sets is called a topology on X .

Remark 2.7. If a set $\kappa \subset X$ can be written in the form $\kappa = X \backslash U$, where U is an open set, then κ is called a closed set.

In optimization problems, we are often interested in whether feasible sets are open or closed.

3. Zero dimensional case

3.1. The roots of polynomials

The quest to solve polynomial equations stands as one of the oldest and most fundamental challenges in mathematics. This has particular significance in optimization theory, where the objective function and constraints often appear as polynomial systems.

The systematic study of polynomial roots accelerated during the 16th century with del Ferro, Tartaglia, and Cardano's solutions to cubic equations, followed by Ferrari's quartic formula [2]. These triumphs revealed an unsettling phenomenon: solutions often required intermediate steps involving square roots of negative numbers, even when the final answers were real. This mystery persisted until Gauss's 1799 doctoral dissertation, where he established the Fundamental Theorem of Algebra:

Theorem 3.1 (Gauss). Every non-constant polynomial with complex coefficients has at least one complex root. Equivalently, a polynomial of degree n has exactly n roots in $\mathbb{C}$, counted with multiplicity.

Example 3.2. The equation below has 3 solutions, namely $1, 1+\sqrt{15}i/2, 1-\sqrt{15}i/2$:

$$f(x) = x^3 - 2x^2 + 5x - 4. \tag{2}$$

Gauss provided four distinct proofs throughout his career, establishing complex numbers as the natural setting for algebraic equations. This profound insight transformed root-finding: rather than

DOI: 10.54254/2753-8818/2026.CH31695

seeking explicit formulas (which are impossible for degree ≥ 5 by Abel–Ruffini [3]), mathematicians could now guarantee the existence of solutions and develop approximation methods.

3.2. The methods of calculating the roots of polynomials

While complex roots complete the algebraic picture, optimization problems usually take place over $\mathbb{R}$. Determining real roots efficiently requires specialized techniques that exploit ordering and continuity, properties absent in $\mathbb{C}$. Three landmark methods demonstrate increasing sophistication in bounding and isolating real roots.

Theorem 3.3 (Descartes). The number of positive real roots of a polynomial, counted with multiplicity, is either equal to the number of sign variations in the coefficient sequence $(a_n, \ldots, a_0)$ or less than it by an even number. Zero coefficients are ignored.

Theorem 3.4 (Budan–Fourier). For a polynomial $f(x)$ of degree n , the number of roots in $(a, b]$ is at most $V_f(a) - V_f(b)$, and differs from this number by an even integer. Here $V_f(c)$ denotes the number of sign changes in the sequence

$$\left(f(c), f'(c), f''(c), \ldots, f^{(n)}(c)\right), \tag{3}$$

and multiple roots are counted once.

Corollary 3.5. Theorem 10 is a special case of Theorem 11, where the interval $(a, b]$ is replaced by $(0, +\infty)$.

Proof. In this case, $V_f(b) = \lim_{x \to +\infty} V_f(x) = 0$, since the sign of the leading coefficient is preserved under differentiation, and for $x \to +\infty$ the leading term determines the sign of $f(x)$.

For $V_f(0)$, if $f = \sum_{i=0}^{n} a_i x^i$, then $f^{(k)}(0) = a_k$, so the sequence $\left(f(0), f'(0), f''(0), \ldots, f^{(n)}(0)\right)$ coincides with $(a_0, \ldots, a_n)$.

Definition 3.6. The Sturm sequence for a square-free polynomial f is defined as

$$f_0 = f, \quad f_1 = f', \quad f_k = -\text{rem}(f_{k-2}, f_{k-1}), \tag{4}$$

where $\text{rem}(a, b)$ denotes the remainder when a is divided by b . The sequence ends when a constant is obtained.

Let $\sigma_f(c)$ denote the number of sign changes in $(f_0(c), \ldots, f_m(c))$, ignoring zeros.

Theorem 3.7 (Sturm). If $a < b$ are not roots of f , then the number of distinct real roots of f in (a, b) is $\sigma_f(a) - \sigma_f(b)$.

Proof. A proof can be found in Sturm's original 1835 paper.

Example 3.8. Suppose $f(x) = x^3 - 2x + 5$. We determine the roots using the above theorems.

By Theorem 3.3: The coefficients of f are $(+, 0, -, +)$, giving two sign changes, which implies either 2 or 0 positive real roots.

By Theorem 3.4: For the interval $(-3, 0]$, the sequence at $x = 0$ is $(5, -2, 0, 6)$, while at $x = -3$ it is $(-16, 25, -18, 6)$. Thus $V_f(-3) - V_f(0) = 3 - 2 = 1$, meaning there is exactly one root in $(-3, 0]$.

By Theorem 3.7:

DOI: 10.54254/2753-8818/2026.CH31695

$$\begin{aligned} f_0 &= x^3 - 2x + 5, \\ f_1 &= 3x^2 - 2, \\ f_2 &= 14/9x - 5, \\ f_3 &= -2361/196 \quad (\text{constant}). \end{aligned} \tag{5}$$

On $[0,1]:\ \sigma_f(0) = 1$ (signs $+,-,-,-$), $\sigma_f(1) = 1$ (signs $+,+,+,-$), giving exactly zero root in $(0,1)$.

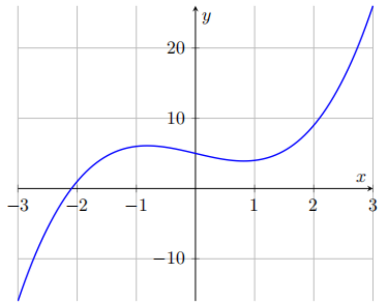

Figure 1. Cubic function $\mathrm{f}(\mathrm{x}) = \mathrm{x}^3 - 2\mathrm{x} + 5$

This figure 1 of $f(x) = x^3 - 2x + 5$ on $[-3,3]$ illustrates the results obtained by the three theorems.

The methods above revolutionized real algebraic geometry and remain foundational for cylindrical algebraic decomposition in optimization.

4. Linear programming to semidefinite programming

In optimization theory, mathematicians usually model real-world situations as mathematical programs that seek to maximize or minimize specific real functions subject to constraints, thereby simplifying complicated real-world problems into calculable forms.

Application 4.1. The Chinese Postman Problem (CPP) is a well-known problem in Graph Theory. It aims to find the shortest closed walk that visits every edge of a graph at least once and returns to the starting point. In undirected graphs, CPP is solvable in polynomial time, while certain directed variants or extensions (e.g., Rural Postman Problem) are NP-hard.

Application 4.2. The Shortest Vector Problem (SVP) is a fundamental computational problem in Lattice Theory. It asks for the shortest non-zero vector in a given lattice, a discrete subgroup of $\mathbb{R}^n$ generated by linearly independent basis vectors. SVP is NP-hard under randomized reductions, making exact solutions computationally intractable for arbitrary lattices. Approximation algorithms such as the Lenstra–Lenstra–Lovász (LLL) algorithm provide polynomial-time methods for finding relatively short vectors. SVP connects to optimization because it can be formulated as a minimization problem over a discrete set, bridging number theory, geometry, and combinatorial optimization, with important applications in lattice-based cryptography.

4.1. Linear programming

The study of linear programming (LP) has a rich historical development. Fourier developed the first systematic method to solve systems of linear inequalities via successive elimination of variables, now called Fourier–Motzkin elimination [4]. Kantorovich and Leontief later studied LP for resource allocation. Others also showed some examples of application such as economics. Dantzig introduced the simplex method, providing a practical (though not polynomial-time) algorithm, transforming optimization from a theoretical concept to an industrial-scale tool [5]. Khachiyan later proved that LP can be solved in polynomial time using the ellipsoid method [6].

Definition 4.3 (Linear Programming). Linear programming problems have the general form:

$$\max_{\mathbf{x}} \mathbf{c}^\top \mathbf{x} \quad \text{subject to} \quad A\mathbf{x} \succcurlyeq \mathbf{b}, \tag{6}$$

where $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{x} \in \mathbb{R}^n$, and $A \in \mathbb{R}^{m \times n}$. The symbol $\succcurlyeq$ represents componentwise inequalities ($\geq$, $\leq$, $=$, etc.).

Commonly used LP solution approaches include:

1. Simplex Method (Dantzig): Traverses the vertices of a polyhedron via pivot operations. Efficient in practice but exponential in the worst case [7].

2. Interior-Point Methods (Karmarkar): Approaches the optimum along interior paths, polynomial-time complexity $O\left(n^{3.5}L\right)$[8].

3. Ellipsoid Algorithm (Khachiyan): The first polynomial-time method ($O\left(n^6L\right)$), theoretically important but rarely used in practice [6].

4.2. Semidefinite Programming

Semidefinite Programming (SDP) emerged in the 1990s as a generalization of LP. Nesterov and Nemirovski extended interior-point methods to SDP [9]. Additionally, other mathematicians expand this theory.

Definition 4.4 (Positive Semidefinite Matrix). A symmetric matrix $X \in \mathbb{R}^{n \times n}$ is positive semidefinite (PSD), denoted $X \succcurlyeq 0$, if $\mathbf{v}^\top X \mathbf{v} \geq 0$ for all $\mathbf{v} \in \mathbb{R}^n$, equivalently, if all eigenvalues of X satisfy $\lambda_i(X) \geq 0$.

Definition 4.5 (Semidefinite Programming). Given symmetric matrices $C, A_i \in \mathbb{S}^n$ and $\mathbf{b} \in \mathbb{R}^m$, solve

$$\begin{aligned} \max_{X \in \mathbb{S}^n} \quad & \langle C, X \rangle = \operatorname{tr}\left(C^\top X\right) \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i, \quad i = 1, \ldots, m, \\ & X \succcurlyeq 0. \end{aligned} \tag{7}$$

Here, $\mathbb{S}^n$ denotes the set of $n \times n$ symmetric matrices. As the figure 2 below shown.

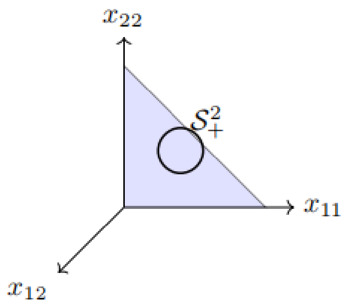

Figure 2. PSD region $\mathscr{S}^2_+$ for 2×2 symmetric matrices

DOI: 10.54254/2753-8818/2026.CH31695

Common SDP solution approaches include:

• Primal-Dual Interior-Point Methods: $O\left(\sqrt{n}\log\left(1/\epsilon\right)\right)$ iterations using Nesterov–Todd directions [9,10].

• First-Order Methods: ADMM with $O\left(1/\epsilon\right)$ complexity for large-scale problems [11].

• Cutting-Plane Methods: For combinatorial optimization with SDP relaxations [12,13].

Some applications of SDP include:

1. Polynomial Optimization: $\min_{\mathbf{x}} p\left(\mathbf{x}\right) \rightarrow \max_{\lambda}\left\{\lambda : p\left(\mathbf{x}\right) - \lambda \text{is a sum-of-squares polynomial}\right\}.$

• Solved via SDP using a monomial basis $Z\left(\mathbf{x}\right)$.

2. Quantum Information: Entanglement detection via the PPT criterion $\rho^{T_B} \succcurlyeq 0$.

3. Robotics: Stability certificates via Lyapunov functions $V\left(\mathbf{x}\right) = \mathbf{x}^{\top} P\mathbf{x}$ with $P \succ 0$ [14,15].

4.3. Polytope, semi-algebraic set, and spectrahedra

Definition 4.6 (Polyhedron and Polytope). A polyhedron $P \subseteq \mathbb{R}^n$ is a convex set defined as the intersection of finitely many closed half-spaces:

$$P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b},\ C\mathbf{x} = \mathbf{d}\}, \tag{8}$$

where $A \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$, $C \in \mathbb{R}^{p \times n}$, and $\mathbf{d} \in \mathbb{R}^p$. If P is bounded, it is called a polytope.

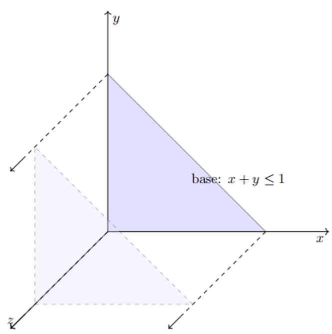

Figure 3. Three-dimensional region with triangular base

The figure 3 above illustrates a polyhedron defined by

$$x \geq 0, y \geq 0, x + y \leq 1, z \geq 0. \tag{9}$$

Definition 4.7 (Semi-Algebraic Set). A set $S \subseteq \mathbb{R}^n$ is called semi-algebraic if it can be expressed as a finite union of sets of the form

$$\{\mathbf{x} \in \mathbb{R}^n : p_1\left(\mathbf{x}\right) = 0, \ldots, p_k\left(\mathbf{x}\right) = 0, q_1\left(\mathbf{x}\right) > 0, \ldots, q_m\left(\mathbf{x}\right) > 0\}, \tag{10}$$

where p_i and q_j are real polynomials in n variables. Equivalently, a semi-algebraic set is any set defined by a finite Boolean combination of polynomial equations and inequalities.

Remark 4.8. An affine real algebraic set is the zero set of a system of polynomials. Every semi-algebraic set can be obtained by projecting a real algebraic set onto selected coordinates. For

DOI: 10.54254/2753-8818/2026.CH31695

example, the semi-algebraic set $x \geq 0$ in $\mathbb{R}$ can be obtained by projecting the algebraic set $\{(x, y) \in \mathbb{R}^2 : xy^2 = 1\}$ onto the x -axis.

Proposition 4.9. The class of semi-algebraic sets is closed under finite unions, finite intersections, and complements.

Proof. This follows directly from the definition of semi-algebraic sets as finite Boolean combinations of polynomial equations and inequalities.

Example 4.10. The set of integers $\mathbb{Z}$ is not a semi-algebraic subset of $\mathbb{R}$. Just like the figure 4 shown.

Lemma 4.11. In any semi-algebraic set $S \subseteq \mathbb{R}^n$, the set of isolated points is finite. An isolated point $\mathbf{x} \in S$ is one for which there exists $\epsilon > 0$ such that

$$B(\mathbf{x}, \epsilon) \cap S = \{\mathbf{x}\}, \tag{11}$$

where $B(\mathbf{x}, \epsilon)$ denotes the open ball centered at $\mathbf{x}$ with radius ϵ .

Proof. It is known that any semi-algebraic set has finitely many connected components, each of which is semi-algebraic. Isolated points correspond to zero-dimensional connected components. Therefore, there can only be finitely many isolated points.

Proof that $\mathbb{Z}$ is not semi-algebraic. Assume, for contradiction, that $\mathbb{Z}$ is semi-algebraic. Every $k \in \mathbb{Z}$ is isolated; for $\epsilon = 1/2$,

$$B(k, 1/2) \cap \mathbb{Z} = \{k\}. \tag{12}$$

Thus, $\mathbb{Z}$ has infinitely many isolated points, contradicting Lemma 4.11. Therefore, $\mathbb{Z}$ cannot be semi-algebraic.

Theorem 4.12. Let $\mathscr{X}$ be a semi-algebraic set in $\mathbb{R}^m$. Let $\mathscr{Y} \subseteq \mathbb{R}^n$ be obtained by projecting $\mathscr{X}$ onto the first n coordinates. Then $\mathscr{Y}$ is semi-algebraic. In short, semi-algebraic sets are stable under projection.

Using this theorem, one can show that a set is not semi-algebraic if its projection onto the x -axis is not semi-algebraic (cf. Lemma 4.11).

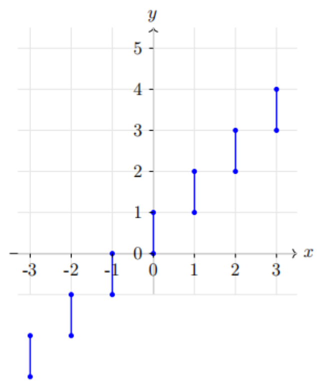

Figure 4. Interger points and line segments

DOI: 10.54254/2753-8818/2026.CH31695

5. Solving a general optimization problem

5.1. Linear programming

Linear programming (LP) can be solved in polynomial time, as proved by Khachiyan [6]. We will not delve into Khachiyan's ellipsoid method; instead, we focus on the theoretical aspect, namely the dual problem.

Definition 5.1 (Dual Problem of LP). For a linear programming problem (Primal)

$$\begin{array}{ll} \text{minimize} & c^{\mathsf{T}}x \\ \text{subject to} & Ax = b, \\ & x \geq 0, \end{array} \tag{13}$$

its dual problem (Dual) is defined as

$$\begin{array}{ll} \text{maximize} & b^{\mathsf{T}}y \\ \text{subject to} & A^{\mathsf{T}}y \leq c. \end{array} \tag{14}$$

Theorem 5.2 (Strong Duality of Linear Programming). Suppose the primal problem (P) and the dual problem (D) are both feasible. Then both problems attain optimal solutions $x^* \in \mathbb{R}^n$, $y^* \in \mathbb{R}^m$, and their optimal values coincide:

$$c^{\mathsf{T}}x^* = b^{\mathsf{T}}y^*. \tag{15}$$

Example 5.3. Consider the LP:

$$\begin{array}{ll} \text{minimize} & 2x_1 + 3x_2 \\ \text{subject to} & x_1 + x_2 \geq 4, \\ & x_1 + 2x_2 \geq 5, \\ & x_1, x_2 \geq 0. \end{array} \tag{16}$$

Its dual is:

$$\begin{array}{ll} \text{maximize} & 4y_1 + 5y_2 \\ \text{subject to} & y_1 + y_2 \leq 2, \\ & y_1 + 2y_2 \leq 3, \\ & y_1, y_2 \geq 0. \end{array} \tag{17}$$

Both problems have optimal value 9.

5.2. Semidefinite programming

Semidefinite programming (SDP) generalizes LP by optimizing over positive semidefinite matrices. The standard primal form is

$$\begin{array}{ll} \text{minimize} & \langle C, X \rangle \\ \text{subject to} & \langle A_i, X \rangle = b_i, \quad i = 1, \ldots, m, \\ & X \succcurlyeq 0, \end{array} \tag{18}$$

where $C, A_i \in \mathbb{S}^n$ ($n \times n$ real symmetric matrices), $b \in \mathbb{R}^m$, $\langle A, B \rangle = \operatorname{tr}\left(A^\top B\right)$ is the Frobenius inner product, and $X \succcurlyeq 0$ denotes positive semidefiniteness.

Definition 5.4 (Dual Problem of SDP). The dual problem (D) associated with (P) is:

$$\begin{array}{ll} \text{maximize} & b^\top y \\ \text{subject to} & \sum_{i=1}^m y_i A_i \preccurlyeq C, \end{array} \tag{19}$$

where $y \in \mathbb{R}^m$, and $\preccurlyeq$ denotes the matrix inequality ($C - \sum y_i A_i \succcurlyeq 0$).

Theorem 5.5 (Weak Duality). For any feasible X and y ,

$$\langle C, X \rangle \geq b^\top y. \tag{20}$$

Equality holds if and only if X and y are optimal.

Theorem 5.6 (Strong Duality of SDP). If either (P) or (D) is strictly feasible (i.e., $\exists X \succ 0$ or $\exists y$ with $C - \sum y_i A_i \succ 0$), then

$$\min_P \left\langle C, X \right\rangle = \max_D b^\top y, \tag{21}$$

and both optima are attained. Otherwise, a duality gap may exist.

Property 5.7 (Computational Aspects).

• SDPs can be solved to ϵ -accuracy in polynomial time via interior-point methods ($O\left(\sqrt{n}\log\left(1/\epsilon\right)\right)$ iterations).

• Exact solutions may be unattainable if the feasible region is empty or unbounded.

• Many NP-hard combinatorial problems (e.g., MAXCUT) admit SDP relaxations.

6. The positivity of polynomials

In optimization problems, it is common to compute the minimum or maximum of a polynomial. This value often yields an optimal solution. Without loss of generality, such problems can be transformed into checking the positivity of a polynomial.

However, even for polynomials without restrictions on the domain, determining non-negativity can be challenging. This question was first considered by Hilbert in 1900.

6.1. Hilbert's 17th problem

In 1900, David Hilbert proposed 23 problems that profoundly influenced 20th-century mathematics. Among them, the 17th problem is central to polynomial optimization and real algebraic geometry.

Problem 6.1 (Hilbert's 17th Problem). Given a multivariate polynomial that is non-negative over $\mathbb{R}^n$, can it be represented as a sum of squares of rational functions?

Historically, Hilbert discussed this with Hermann Minkowski. While Minkowski conjectured that some non-negative polynomials cannot be expressed as sums of squares of polynomials, Hilbert

DOI: 10.54254/2753-8818/2026.CH31695

maintained a more optimistic view for their representability. Hilbert himself established that:

• For univariate polynomials, non-negativity is equivalent to being a sum of squares.

• For bivariate polynomials of degree 4, he constructed the first example of a non-negative polynomial not expressible as a sum of squares of polynomials.

The complete resolution came in 1927 through Emil Artin [16]:

Theorem 6.2 (Artin's Solution to Hilbert's 17th Problem). If $f \in \mathbb{R}[x_1, \ldots, x_n]$ is non-negative on $\mathbb{R}^n$, then there exist rational functions $g_1, \ldots, g_k \in \mathbb{R}(x_1, \ldots, x_n)$ such that $f = g_1^2 + \cdots + g_k^2$. Equivalently, there exist polynomials $p_1, \ldots, p_k, q \in \mathbb{R}[x_1, \ldots, x_n]$ with $q \neq 0$ such that

$$f = \frac{p_1^2 + \cdots + p_k^2}{q^2}. \tag{22}$$

Despite this theoretical resolution, explicit examples of non-negative polynomials that are not sums of squares of polynomials were discovered later. The first explicit example is due to Theodore Motzkin [17]:

Example 6.3 (Motzkin Polynomial). The polynomial

$$M(x, y) = x^4y^2 + x^2y^4 - 3x^2y^2 + 1 \tag{23}$$

is non-negative on $\mathbb{R}^2$ but cannot be expressed as a sum of squares of polynomials in $\mathbb{R}[x, y]$.

The proof involves algebraic geometry techniques, including Newton polytopes and homogeneous components. The Motzkin polynomial illustrates the distinction between non-negativity and sums-of-squares representability.

6.2. Convex cones and SOS polynomials

Definition 6.4 A set $C \subseteq \mathbb{R}^n$ is a convex cone if:

1. Convexity: $\forall x, y \in C, \forall \theta \in [0,1],\ \theta x + (1-\theta)y \in C$.
2. Cone property: $\forall x \in C, \forall \lambda \geq 0,\ \lambda x \in C$.

Lemma 6.5. The set of sum-of-squares (SOS) polynomials of degree at most $2d$ in n variables,

$$\mathrm{SOS}_{\leq 2d}[x_1, \ldots, x_n], \tag{24}$$

is a convex cone.

Proof. Let $p = \sum_{i=1}^{k} f_i^2$ and $q = \sum_{j=1}^{m} g_j^2$ be SOS polynomials.

Cone property: For $\lambda \geq 0$,

$$\lambda p = \sum_{i=1}^{k} \left(\sqrt{\lambda} f_i\right)^2 \in \mathrm{SOS}_{\leq 2d}[x_1, \ldots, x_n]. \tag{25}$$

Convexity: For $\theta \in [0,1]$,

$$\theta p + (1-\theta)q = \sum_{i=1}^{k} \left(\sqrt{\theta} f_i\right)^2 + \sum_{j=1}^{m} \left(\sqrt{1-\theta} g_j\right)^2 \in \mathrm{SOS}_{\leq 2d}[x_1, \ldots, x_n]. \tag{26}$$

Lemma 6.6. The set of non-negative polynomials of degree at most $2d$,

Proceedings of CONF-CIAP 2026 Symposium: Applied Mathematics and Statistics
DOI: 10.54254/2753-8818/2026.CH31695

$$\mathbb{R}_{\leq 2d}[x_1,\ldots,x_n]^+ = \{f : f(\mathbf{x}) \geq 0\ \forall \mathbf{x} \in \mathbb{R}^n\}, \tag{27}$$

is a convex cone.

Proof. Let $p, q \in \mathbb{R}_{\leq 2d}[x_1,\ldots,x_n]^+$. Then for $\lambda \geq 0$ and $\theta \in [0,1]$, we have

$$(\lambda p)(\mathbf{x}) \geq 0, \quad [\theta p + (1-\theta)q](\mathbf{x}) \geq 0. \tag{28}$$

Remark 6.7. Polynomials can be represented via Gram matrices. Let $z(\mathbf{x})$ be the vector of monomials of degree at most d . Then for $f \in \mathbb{R}_{\leq 2d}[x_1,\ldots,x_n]$,

$$f(\mathbf{x}) = z(\mathbf{x})^\top Q z(\mathbf{x}), \tag{29}$$

where Q is symmetric, called the Gram matrix of f .

Theorem 6.8. A polynomial $f \in \mathbb{R}_{\leq 2d}[x_1,\ldots,x_n]$ is SOS if and only if there exists $Q \succcurlyeq 0$ such that

$$f(\mathbf{x}) = z(\mathbf{x})^\top Q z(\mathbf{x}). \tag{30}$$

Lemma 6.9. The SOS cone $\mathrm{SOS}_{\leq 2d}[x_1,\ldots,x_n]$ is closed.

Proof. The mapping $f \mapsto Q$ is linear. The set of PSD matrices is closed, so the preimage under a continuous linear map is closed.

6.3. Motzkin polynomial and perturbations

Proposition 6.10. The Motzkin polynomial

$$M(x,y) = x^4y^2 + x^2y^4 - 3x^2y^2 + 1 \tag{31}$$

is non-negative on $\mathbb{R}^2$ but not SOS.

Proof. Non-negativity: by AM-GM,

$$\frac{x^4y^2+x^2y^4+1}{3} \geq \sqrt[3]{x^6y^6} = x^2y^2. \tag{32}$$

Non-SOS: any Gram matrix Q for M has negative eigenvalues (submatrix for x^2y and xy^2 terms has eigenvalues $\pm 3/2$), so $M \notin \mathrm{SOS}$.

Theorem 6.11. Let $g \in \mathbb{R}[x_1,x_2]$ be non-negative. For sufficiently small $\epsilon > 0$,

$$M_\epsilon(x,y) = M(x,y) + \epsilon g(x,y) \tag{33}$$

is non-negative but not SOS.

Proof. Non-negativity is preserved since $M, g \geq 0$.

For SOS: the Gram matrix mapping is linear and PSD matrices form a closed set. Since M has a non-PSD Gram matrix, for small ϵ , M_ϵ also has a non-PSD Gram matrix.

Remark 6.12. The Motzkin polynomial lies on the boundary of the SOS cone. Theorem 45 shows the SOS cone is strictly contained in the non-negative cone:

DOI: 10.54254/2753-8818/2026.CH31695

$$\mathrm{SOS}_{\leq 2d}[x_1, \ldots, x_n] \subsetneq \mathbb{R}_{\leq 2d}[x_1, \ldots, x_n]^+. \tag{34}$$

References

[1] G. Strang, Introduction to linear algebra. SIAM, 2022.
[2] G. Cardano, Ars magna or the rules of algebra. New York: Dover Publications, 1993.
[3] M. Artin, Algebra, 2nd ed. Pearson, 2014.
[4] J.-B. J. Fourier, “Histoire de l’académie, partie mathématique, ” Mémoires de l’Académie des sciences de l’Institut de France, vol. 7, pp. xlvii–lv, 1827.
[5] G. B. Dantzig, Linear programming and extensions. Princeton: Princeton University Press, 1963.
[6] L. G. Khachiyan, “A polynomial algorithm in linear programming, ” Doklady Akademii Nauk SSSR, vol. 244, no. 5, pp. 1093–1096, 1979.
[7] G. B. Dantzig, “Maximization of a linear function of variables subject to linear inequalities, ” Activity Analysis of Production and Allocation, pp. 339–347, 1951.
[8] N. Karmarkar, “A new polynomial-time algorithm for linear programming, ” Combinatorica, vol. 4, no. 4, pp. 373–395, 1984.
[9] Y. Nesterov and A. Nemirovskii, Interior-point polynomial algorithms in convex programming, vol. 13. SIAM, 1994.
[10] S. J. Wright, Primal-dual interior-point methods. SIAM, 1997.
[11] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers, ” Foundations and Trends in Machine Learning, vol. 3, no. 1, pp. 1–122, 2011.
[12] M. Grötschel, L. Lovász, and A. Schrijver, Geometric algorithms and combinatorial optimization, vol. 2. Springer-Verlag, 1988.
[13] C. Helmberg, “Semidefinite programming for combinatorial optimization, ” PhD thesis, Technische Universität Berlin, 2000.
[14] J.-J. E. Slotine and W. Li, Applied nonlinear control. Prentice Hall, 1991.
[15] H. K. Khalil, Nonlinear systems, 3rd ed. Prentice Hall, 2002.
[16] E. Artin, “Über die zerlegung definiter funktionen in quadrate, ” Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg, vol. 5, pp. 110–115, 1927.
[17] T. S. Motzkin, “The arithmetic-geometric inequality, ” in Proc. Symposium on inequalities, O. Shisha, Ed., New York: Academic Press, 1967, pp. 205–224.