Appendix A
Background Material

Grigoriy Blekherman, Pablo A. Parrilo, and Rekha R. Thomas

The appendix consists of four parts: matrices and quadratic forms, convex optimization, convex geometry, and algebraic geometry. The material in this appendix is mostly standard and as such is presented for the convenience of the reader in a compact form.

A.1 Matrices and Quadratic Forms

We present here a few basic facts about linear algebra, symmetric matrices, and quadratic forms. There are many excellent references on the topic, including [11] and [15], among others.

A matrix $A \in \mathbb{R}^{n\times n}$ is *symmetric* if $a_{ij} = a_{ji}$ for $i, j = 1, \ldots, n$. The set of symmetric matrices is denoted as $\mathcal{S}^n$ and is a real vector space of dimension $\binom{n+1}{2} = \frac{1}{2}(n+1)n$. Real quadratic forms can always be represented in terms of symmetric matrices, i.e., $q(x) = \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_ix_j = x^TAx$, where $a_{ij} = a_{ji}$. We often identify a symmetric matrix with the corresponding quadratic form.

The *characteristic polynomial* of a matrix $A \in \mathcal{S}^n$ is $p_A(\lambda) := \det(\lambda I - A) = \lambda^n + \sum_{k=0}^{n-1} p_k\lambda^k = \prod_{k=1}^n(\lambda - \lambda_k)$, where λ_k are the *eigenvalues* of A. Given a subset $S \subseteq \{1, \ldots, n\}$, let A_S be the submatrix of A whose rows and columns are indexed by S. The *principal minor* of A corresponding to the subset S is the determinant of A_S. If S has the form $\{1, 2, \ldots, k\}$, then the corresponding minor is called a *leading principal minor*. It can be shown that the coefficient p_k of the characteristic polynomial is equal (up to sign) to the sum of all the principal minors of size $n-k$, i.e., $p_k = (-1)^{n-k}\sum_{S:|S|=n-k} \det A_S$. Notice that, in particular, $p_{n-1} = -\mathrm{Tr}\, A$ and $p_0 = (-1)^n \det A$.

A.1.1 Positive Semidefinite Matrices

If the quadratic form $x^T Ax$ takes only nonnegative values, we say that the matrix A is *positive semidefinite*. Similarly, if it takes only positive values (except at the origin, where it necessarily vanishes), then A is *positive definite*. There are several equivalent conditions for a matrix to be positive (semi)definite:

Proposition A.1. *Let $A \in \mathcal{S}^n$ be a symmetric matrix. The following statements are equivalent:*

1. *The matrix A is positive semidefinite ($A \succeq 0$).*
2. *For all $x \in \mathbb{R}^n$, $x^T Ax \geq 0$.*
3. *All eigenvalues of A are nonnegative.*
4. *All $2^n - 1$ principal minors of A are nonnegative.*
5. *The coefficients of $p_A(\lambda)$ weakly alternate in sign, i.e., $(-1)^{n-k} p_k \geq 0$ for $k = 0, \ldots, n-1$.*
6. *There exists a factorization $A = BB^T$, where $B \in \mathbb{R}^{n \times r}$ and r is the rank of A.*

For the definite case, there are similar characterizations:

Proposition A.2. *Let $A \in \mathcal{S}^n$ be a symmetric matrix. The following statements are equivalent:*

1. *The matrix A is positive definite ($A \succ 0$).*
2. *For all nonzero $x \in \mathbb{R}^n$, $x^T Ax > 0$.*
3. *All eigenvalues of A are strictly positive.*
4. *All n leading principal minors of A are strictly positive.*
5. *The coefficients of $p_A(\lambda)$ alternate in sign, i.e., $(-1)^{n-k} p_k > 0$ for $k = 0, \ldots, n-1$.*
6. *There exists a factorization $A = BB^T$, with B square and nonsingular.*

The set of positive semidefinite matrices is denoted as $\mathcal{S}^n_+$, and its interior (the set of positive definite matrices) as $\mathcal{S}^n_{++}$. The set $\mathcal{S}^n_+$ is invariant under nonsingular *congruence transformations*; i.e., if T is nonsingular, $A \succeq 0 \Leftrightarrow T^T AT \succeq 0$. The same statement holds for its interior, i.e., $A \succ 0 \Leftrightarrow T^T AT \succ 0$. For additional facts about the geometry of the set of positive semidefinite matrices, see Section A.3.5.

A.1.2 Matrix Factorizations

For a symmetric matrix A, there are several *matrix factorizations* that can be used to determine or certify properties of A; see, e.g., [11] for theoretical background and [9] for computational aspects. Among the most important matrix factorizations, we have the following.

Eigenvalue decomposition. Since A is symmetric, the eigenspaces corresponding to distinct eigenvalues are mutually orthogonal, and thus one can choose an orthonormal basis of eigenvectors. As a consequence, the matrix A is diagonalizable and there is always a decomposition

$$A = V\Lambda V^T, \qquad \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n),$$

where the matrix V is orthogonal ($VV^T = V^TV = I$). If A is positive semidefinite, we have $\lambda_i \geq 0$ for $i = 1, \ldots, n$.

Cholesky decomposition. A positive semidefinite matrix A can be decomposed as

$$A = LL^T,$$

where L is a lower triangular matrix (i.e., $L_{ij} = 0$ for $j > i$). This is known as the *Cholesky decomposition* of the matrix A and can be obtained by solving the identity above column by column (or row by row). The Cholesky decomposition can be computed in $O(n^3)$ operations (in the dense case, faster if the matrix is sparse). Notice that, as opposed to eigenvalue methods, no iterative methods are required. This decomposition plays an important role in numerical algorithms for semidefinite programming.

LDL^T decomposition. This is a decomposition of the form

$$A = LDL^T,$$

where the matrix D is diagonal with nonnegative entries, and L is lower triangular with normalized diagonal entries $L_{ii} = 1$. It should be clear that this can be directly obtained from the Cholesky decomposition, by suitably normalizing its diagonal entries. The importance of the LDL^T decomposition is that, in contrast to the other two factorizations discussed above, it is a *rational* decomposition; i.e., if the matrix A is rational then all numbers that appear in the decomposition are rational (and also, polynomially sized).

Two distinct factorizations of the same positive semidefinite matrix can always be related through a suitable orthogonal transformation. The following result makes this precise.

Theorem A.3. *Let $A \in \mathcal{S}^n$ be a positive semidefinite symmetric matrix, with $A = FF^T$ and $A = GG^T$, where $F, G \in \mathbb{R}^{n\times p}$. Then, there exists a matrix $U \in \mathbb{R}^{p\times p}$ such that $F = GU$ and U is orthogonal (i.e., such that $UU^T = I$ and $U^TU = I$).*

A.1.3 Inertia and Signature

Definition A.4. *Consider a symmetric matrix A. The* inertia *of A, denoted $\mathcal{I}(A)$, is the integer triple (n_-, n_0, n_+), where n_-, n_0, n_+ are the number of negative, zero, and positive eigenvalues, respectively. The* signature *of A is equal to the number*

of positive eigenvalues minus the number of negative eigenvalues, i.e., the integer $n_+ - n_-$.

Notice that, with the notation above, the rank of A is equal to $n_+ + n_-$. A symmetric positive definite $n \times n$ matrix has inertia $(0, 0, n)$, while a positive semidefinite one has $(0, k, n-k)$ for some $k \geq 0$.

The inertia is an important invariant of a quadratic form, since it holds that $\mathcal{I}(A) = \mathcal{I}(TAT^T)$, where T is nonsingular. This invariance of the inertia of a matrix under congruence transformations is known as *Sylvester's law of inertia*; see, for instance, [11]. This invariance makes it possible to efficiently compute the inertia of a matrix from its LDL^T decomposition, since in this case $\mathcal{I}(A) = \mathcal{I}(D)$, and the inertia of a diagonal matrix is trivial to compute.

A.1.4 Schur Complements

Given a block-partitioned matrix

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix},$$

where A is square and invertible, the *Schur complement* of A is the matrix $C - B^T A^{-1} B$. Similarly, if C is square and invertible, its Schur complement is the matrix $A - BC^{-1}B^T$. Schur complements appear in many areas, including among others convex optimization (partial minimization of quadratic functions), probability and statistics (conditioning and marginalization of multivariate Gaussians), and algorithms (block matrix inversion). For several applications and generalizations, see, for instance, the classical survey [6].

Many of the properties of the Schur complement follow quite directly from the two factorizations:

$$\begin{aligned} \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} &= \begin{bmatrix} I & 0 \\ B^T A^{-1} & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & C - B^T A^{-1} B \end{bmatrix} \begin{bmatrix} I & A^{-1}B \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} I & BC^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} A - BC^{-1}B^T & 0 \\ 0 & C \end{bmatrix} \begin{bmatrix} I & 0 \\ C^{-1}B^T & I \end{bmatrix}. \end{aligned}$$

Since the factorizations above are congruence transformations, this implies that the following conditions are equivalent:

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \succ 0 \quad \Leftrightarrow \quad \begin{cases} A \succ 0, \\ C - B^T A^{-1} B \succ 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} C \succ 0, \\ A - BC^{-1}B^T \succ 0. \end{cases}$$

A.2 Convex Optimization

In this section we describe the basic elements of optimization theory, with an emphasis on convexity. For additional background, complete statements, and proofs, we refer the reader to the works [2, 3, 5].

A.2.1 Convexity and Hessians

A set $S \subset \mathbb{R}^n$ is a *convex set* if $x, y \in S$ implies $\lambda x + (1-\lambda)y \in S$ for all $0 \le \lambda \le 1$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is a *convex function* if $f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda)f(y)$ for all $0 \le \lambda \le 1$ and $x, y \in \mathbb{R}^n$. A function f is convex if and only if its *epigraph* $\{(x,t) \in \mathbb{R}^n \times \mathbb{R} : f(x) \le t\}$ is a convex set. A function f is *concave* if $-f$ is convex. When a function is differentiable there are several equivalent characterizations of convexity, in terms of the gradient $\nabla f(x)$ or the Hessian $\nabla^2 f(x)$:

Lemma A.5. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a twice differentiable function. The following propositions are equivalent:*

(i) *The function f is convex, i.e.,*

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda)f(y) \quad \textit{for all} \quad 0 \le \lambda \le 1, \quad x, y \in \mathbb{R}^n.$$

(ii) *The* first-order *convexity condition holds:*

$$f(y) \ge f(x) + (\nabla f(x))^T (y - x), \quad \textit{for all} \quad x, y \in \mathbb{R}^n.$$

(iii) *The* second-order *convexity condition holds:*

$$\nabla^2 f(x) \succeq 0, \quad \textit{for all} \quad x \in \mathbb{R}^n,$$

i.e., the Hessian is positive semidefinite everywhere.

A.2.2 Minimax Theorem

Given a function $f : S \times T \to \mathbb{R}$, the following inequality always hold:

$$\max_{t \in T} \min_{s \in S} f(s,t) \le \min_{s \in S} \max_{t \in T} f(s,t). \tag{A.1}$$

If the maxima or minima in (A.1) are not attained, the inequality is still true by replacing "max" and "min" with "sup" and "inf," respectively.

It is of interest to understand situations under which (A.1) holds with equality. The following is a well-known condition for this.

Theorem A.6 (minimax theorem). *Let $S \subset \mathbb{R}^n$ and $T \subset \mathbb{R}^m$ be compact convex sets, and $f : S \times T \to \mathbb{R}$ be a continuous function that is convex in its first argument and concave in the second. Then*

$$\max_{t \in T} \min_{s \in S} f(s,t) = \min_{s \in S} \max_{t \in T} f(s,t).$$

A special case of this theorem, used in game theory to prove the existence of equilibria for zero-sum games, is when S and T are standard unit simplices and the function $f(s,t)$ is a bilinear form.

A.2.3 Lagrangian Duality

Consider a nonlinear optimization problem:

$$\begin{array}{ll} \underset{x\in\mathbb{R}^n}{\text{minimize}} & f(x) \\ \text{subject to} & g_i(x) \le 0, \quad i = 1,\dots,m, \\ & h_j(x) = 0, \quad j = 1,\dots,p, \end{array} \tag{A.2}$$

and let $u^\star$ be its optimal value. Define the *Lagrangian* associated with the optimization problem (A.2) as

$$\begin{array}{rcl} L : \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^p & \to & \mathbb{R}^n, \\ (x,\lambda,\mu) & \mapsto & f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{j=1}^p \mu_j h_j(x). \end{array}$$

The *Lagrange dual function* is defined as

$$\phi(\lambda,\mu) := \min_{x\in\mathbb{R}^n} \quad L(x,\lambda,\mu),$$

Maximizing this function over the dual variables (μ,λ) yields

$$v^\star := \max_{\mu\in\mathbb{R}^p \text{ and } \lambda\ge 0} \phi(\lambda,\mu).$$

Applying the minimax inequality (A.1), we see that this is a lower bound on the value of the original optimization problem:

$$v^\star \le \min_{x\in\mathbb{R}^n} \max_{\mu\in\mathbb{R}^p \text{ and } \lambda\ge 0} L(x,\lambda,\mu) = u^\star.$$

If the functions f, g_i are convex and h_i are affine, then the Lagrangian is convex in x and concave in (λ,μ). To ensure strong duality (i.e., equality in the expression above), compactness or other *constraint qualifications* are needed. An often used condition is the *Slater constraint qualification*: there exists a strictly feasible point, i.e., a point $z^\star \in \mathbb{R}^n$ such that $g_i(z^\star) < 0$ for all $i = 1,\dots,m$ and $h_j(z^\star) = 0$ for all $j = 1,\dots,p$. Under this condition, strong duality always holds.

Theorem A.7. *Consider the optimization problem* (A.2)*, where f, g_i are convex and h_i are affine. Assume Slater's constraint qualification holds. Then the optimal value of the primal is the same as the optimal value of the dual, i.e., $v^\star = u^\star$.*

A.2.4 KKT Optimality Conditions

Consider the nonlinear optimization problem in (A.2). The *Karush–Kuhn–Tucker* (KKT) optimality conditions are

$$\begin{array}{rrcll} \nabla f\Big|_{x^\star} + \displaystyle\sum_{i=1}^m \lambda_i^\star \cdot \nabla g_i\Big|_{x^\star} + \sum_{j=1}^p \mu_j^\star \cdot \nabla h_j\Big|_{x^\star} & & = & 0, & \\ \text{Primal feasibility:} & g_i(x^\star) & \le & 0 & \text{for } i = 1,\dots,m, \\ & h_j(x^\star) & = & 0 & \text{for } j = 1,\dots,p, \\ \text{Dual feasibility:} & \lambda_i^\star & \ge & 0 & \text{for } i = 1,\dots,m, \\ \text{Complementary slackness:} & \lambda_i^\star \cdot g_i(x^\star) & = & 0 & \text{for } i = 1,\dots,m. \end{array} \tag{A.3}$$

Under certain constraint qualifications (e.g., the ones in the theorem below), the KKT conditions are necessary for local optimality.

Theorem A.8. *Assume any of the following* constraint qualifications *hold:*

- *The gradients of the constraints* $\{\nabla g_1(x^\star), \ldots, \nabla g_m(x^\star), \nabla h_1(x^\star), \ldots, \nabla h_p(x^\star)\}$ *are linearly independent.*
- *There exists a strictly feasible point (Slater constraint qualification), i.e., a point* $z^\star \in \mathbb{R}^n$ *such that* $g_i(z^\star) < 0$ *for all* $i = 1, \ldots, m$ *and* $h_j(z^\star) = 0$ *for all* $j = 1, \ldots, p$.
- *All constraints* $g_i(x)$, $h_i(x)$ *are affine functions.*

Then, at every local minimum $x^\star$ *of* (A.2) *the KKT conditions* (A.3) *hold.*

On the other hand, for *convex optimization problems*, i.e., if all functions f, g_i are convex and h_i are affine, then the KKT conditions are *sufficient* for local (and thus global) optimality:

Theorem A.9. *Let* (A.2) *be a convex optimization problem and* $x^\star$ *be a point that satisfies the KKT conditions* (A.3)*. Then* $x^\star$ *is a global minimum.*

A.3 Convex Geometry

We give a summary of standard properties of convex sets and the cone of positive semidefinite matrices. We refer the reader to [2, 13, 14] for more background and proofs.

A.3.1 Basic Facts

Recall that a subset K of $\mathbb{R}^n$ is called convex if for all $x, y \in K$ we have $\lambda x + (1-\lambda)$ $y \in K$ for all real $0 \le \lambda \le 1$.

For vectors $x_1, \ldots, x_k \in \mathbb{R}^n$ a linear combination $\lambda_1 x_1 + \cdots + \lambda_k x_k$ is called a *convex combination* if $\lambda_i \ge 0$ for $1 \le i \le k$ and $\lambda_1 + \cdots + \lambda_k = 1$. A linear combination is called a *conic combination* if we require only that $\lambda_i \ge 0$ for $1 \le i \le k$. Equivalently, a subset K of $\mathbb{R}^n$ is convex if it is closed under taking convex combinations, and K is *convex cone* if it is closed under taking conic combinations. Equivalently, a convex cone is a convex set that is also closed under multiplication by nonnegative scalars.

Let $S \subseteq \mathbb{R}^n$ be an arbitrary subset. The convex hull, $\mathrm{conv}(S)$, of S is the smallest convex set containing S. Equivalently $\mathrm{conv}(S)$ is the set of all convex combinations of points in S:

$$\mathrm{conv}(S) = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{c} x = \lambda_1 y_1 + \cdots + \lambda_k y_k \quad \text{for some} \quad y_1, \ldots, y_k \in S \\ \lambda_i \ge 0, \quad \lambda_1 + \cdots + \lambda_k = 1 \end{array} \right\}.$$

The conic hull, $\text{cone}(S)$, of S is the set of all conic combinations of the points in S:

$$\text{cone}(S) = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{c} x = \lambda_1 y_1 + \cdots + \lambda_k y_k \quad \text{for some} \quad y_1, \ldots, y_k \in S \\ \lambda_i \geq 0 \end{array} \right\}.$$

The set $\text{cone}(S)$ is the smallest convex cone containing S.

A priori it is not clear how large, in terms of the number of points, the convex combinations of points in S need to be to write down a point in $\text{conv}(S)$. Carathéodory's Theorem provides an upper bound.

Theorem A.10. *Let S be a subset of $\mathbb{R}^n$. Then any point in the convex hull of S can be written as a convex combination of at most $n+1$ points in S.*

A set defined by finitely many linear inequalities is called a *polyhedron*. The convex hull of finitely many points in $\mathbb{R}^n$ is called a *polytope*, and the conic hull of finitely many points in $\mathbb{R}^n$ is called a *polyhedral cone*. We then have the following theorem.

Theorem A.11. *A bounded polyhedron is a polytope.*

Convex sets possess a well-defined notion of dimension. Let $K \subseteq \mathbb{R}^n$ be a convex set and let $\text{Aff}(K)$ be its *affine hull*, i.e., the affine linear subspace spanned by K. Then K has interior, as a subset of $\text{Aff}(K)$ and $\dim K = \dim \text{Aff}(K)$. The interior of K as a subset of $\text{Aff}\, K$ is called the *relative interior* of K, and the boundary of K as a subset of $\text{Aff}\, K$ is called the *relative boundary* of K. A compact convex set that is full dimensional in $\mathbb{R}^n$ is a called a *convex body*.

Let $K \subset \mathbb{R}^n$ be a closed convex set. A subset $F \subseteq K$ is called a *face* of K if for all $x, y \in K$ and any $0 \leq \lambda \leq 1$, if we have $\lambda x + (1-\lambda)y \in F$, then $x, y \in F$. A face F is called a *proper face* of K if F is a nonempty proper subset of K. It is easy to see that a proper face F does not contain any points in the relative interior of K and therefore it is a subset of the relative boundary of K. The intersection of any two faces of K is a face of K.

A face F of K is called *exposed* if there exists an affine hyperplane H in $\mathbb{R}^n$ such that $F = K \cap H$. The hyperplane H divides $\mathbb{R}^n$ into two half spaces, and it follows that $K \setminus F$ must lie in one of the two open half spaces. Equivalently F is an exposed face of K if there exists an affine linear functional $\ell : \mathbb{R}^n \to \mathbb{R}$ such that $\ell(x) \geq 0$ for all $x \in K$ and $\ell(s) = 0$ for all $s \in F$.

A point $x \in K$ is called an *extreme point* of K if x is a face of K; i.e., if $x = \lambda y_1 + (1-\lambda) y_2$ for some $y_1, y_2 \in K$ and $0 \leq \lambda \leq 1$, then $y_1 = y_2 = x$. A point x in a convex cone C is said to span an *extreme ray* of C if the ray spanned by x is a face of C; i.e., if $x = \lambda_1 y_1 + \lambda_2 y_2$ for some $y_1, y_2 \in K$ and $\lambda_1, \lambda_2 \geq 0$, then y_1 and y_2 lie on the ray spanned by x. The following is the finite-dimensional Krein–Milman theorem. It is also known as Minkowski's theorem.

Theorem A.12. *Let $K \subset \mathbb{R}^n$ be a compact convex set. Then K is the convex hull of its extreme points.*

Faces of a compact convex set K ordered by inclusion form a partially ordered set $\mathcal{F}(K)$. The poset $\mathcal{F}(K)$ is a *lattice*, where the meet operation is intersection: $F_1 \vee F_2 = F_1 \cap F_2$ and the join operation is the intersection of all faces containing F_1 and F_2: $F_1 \wedge F_2 = \bigcap_{F \supseteq F_1, F_2} F$.

It follows from the Krein–Milman theorem that the minimal proper faces in $\mathcal{F}(K)$ are the extreme points of K, and it will follow from separation theorems presented below that the maximal proper faces in $\mathcal{F}(K)$ are exposed. We note that the maximal proper faces of $\mathcal{F}(K)$ do not have to have the same dimension. In particular for $K \subset \mathbb{R}^n$, the dimension of all the maximal proper faces can be strictly smaller than $n-1$. This is the case for the cone of positive semidefinite matrices $\mathcal{S}^n_+$ as we will see below.

A.3.2 Cone Decomposition

Let K_1, K_2 be convex subsets of $\mathbb{R}^n$. Define $K_1 + K_2$ as the set of all sums of points from K_1 and K_2:

$$K_1 + K_2 = \{x \in \mathbb{R}^n \mid x = x_1 + x_2,\, x_1 \in K_1,\, x_2 \in K_2\}.$$

This operation is called *Minkowski addition*, and the set $K_1 + K_2$ is also convex.

A closed convex cone $C \subset \mathbb{R}^n$ is called *pointed* if C does not contain straight lines. A cone that is closed, full-dimensional in $\mathbb{R}^n$, and pointed is called a *proper* cone. The following theorem shows that a nonpointed cone can always be decomposed into a pointed cone and a subspace.

Theorem A.13. *Let $C \subset \mathbb{R}^n$ be a closed convex cone. Then C is the Minkowski sum of a pointed cone P and a linear subspace L:*

$$C = P + L.$$

This allows us to concentrate on pointed convex cones. Now we formulate the analogue of the Krein–Milman theorem for pointed cones.

Theorem A.14. *Let C be a closed pointed cone in $\mathbb{R}^n$. Then C is the conic hull of the points spanning its extreme rays.*

There is also decomposition theorem for polyhedra, called the Minkowski–Weyl theorem.

Theorem A.15. *Every polyhedron is a Minkowski sum of a polytope and a polyhedral cone.*

A.3.3 Separation Theorems

An important property of a convex set is that we can *certify* when a point is not in the set. This is usually done via a separation theorem. Let H be an affine hyperplane in $\mathbb{R}^n$. Then H divides $\mathbb{R}^n$ into two half spaces. We will use H_+ and H_- to denote the open half spaces and $\bar{H}_+$ and $\bar{H}_-$ to denote the closed half spaces.

We say that H *separates* two sets K_1 and K_2 if K_1 and K_2 belong to different closed half spaces $\bar{H}_+$ and $\bar{H}_-$. We say that H *strictly separates* K_1 and K_2 if they belong to different open subspaces H_+ and H_-.

Equivalently we can think of H as the zero set of an affine linear functional $\ell : \mathbb{R}^n \to R$. Then ℓ separates K_1 and K_2 if $\ell(x) \geq 0$ for all $x \in K_1$ and $\ell(x) \leq 0$ for all $x \in K_2$. Similarly ℓ strictly separates K_1 and K_2 if $\ell(x) > 0$ for all $x \in K_1$ and $\ell(x) < 0$ for all $x \in K_2$.

Now we state our most general separation theorem.

Theorem A.16. *Let K_1 and K_2 be convex subsets of $\mathbb{R}^n$ such that $K_1 \cap K_2 = \emptyset$. Then there exists an affine hyperplane H that separates K_1 and K_2.*

We observe that it follows from Theorem A.16 that every face of a convex set K is contained in an exposed face of K.

We will often be interested in strict separation, in which case we need to make further assumptions on K_1 and K_2.

Theorem A.17. *Let K_1 and K_2 be disjoint convex subsets of $\mathbb{R}^n$ and suppose that K_1 is compact and K_2 is closed. Then there exists an affine hyperplane H strictly separating K_1 and K_2.*

Theorem A.17 is often applied when K_1 is a single point. Separation theorems lead to certificates of not belonging to a convex set. Combined with notions of polarity explained below this leads to *theorems of the alternative*.

We need to adjust Theorems A.16 and A.17 to the setting of cones, since, for example, all cones contain the origin and are never disjoint. Also, any hyperplane separating two cones C_1 and C_2 must be linear. We will say that a linear functional $\ell : \mathbb{R}^n \to R$ separates C_1 and C_2 if $\ell(x) \geq 0$ for all $x \in C_1$ and $\ell(x) \leq 0$ for all $x \in C_2$. Similarly ℓ strictly separates C_1 and C_2 if $\ell(x) > 0$ for all nonzero $x \in C_1$ and $\ell(x) < 0$ for all nonzero $x \in C_2$. Then we have the following theorem.

Theorem A.18. *Let C_1 and C_2 be pointed closed convex cones in $\mathbb{R}^n$ such that $C_1 \cap C_2 = 0$. Then there exists a linear functional $\ell : \mathbb{R}^n \to \mathbb{R}$ strictly separating C_1 and C_2.*

A.3.4 Polarity and Duality

We can view a compact convex set K as the convex hull of its extreme points, but we can also view it as being cut out by linear inequalities. The set of affine linear inequalities defining K is a convex object itself, and this leads to very fruitful notions of polarity and duality in convex geometry.

Let $\langle\ ,\ \rangle$ be an inner product on $\mathbb{R}^n$. Let $K \subset \mathbb{R}^n$ be a convex body with origin in its interior. Define the *polar* body K° as follows:

$$K^\circ = \{x \in \mathbb{R}^n \ \mid \ \langle x, y\rangle \leq 1 \quad \text{for all} \quad y \in K\}.$$

The polar body encodes all the affine linear defining inequalities of K. It is easy to see that K° is also a convex body with origin in its interior. Moreover $x \in K^\circ$ is

on the boundary of K° if and only if $\langle x, y\rangle = 1$ for some $y \in K$. Polarity reverses inclusion: if K_1 and K_2 are convex bodies and

$$\text{if} \quad K_1 \subseteq K_2, \quad \text{then} \quad K_2^\circ \subseteq K_1^\circ.$$

First we observe that polarity is an involution on convex bodies with origin in the interior.

Theorem A.19 (biduality theorem). *Let K be a convex body with origin in its interior. Then*

$$(K^\circ)^\circ = K.$$

We now note that extreme points of K° define maximal proper faces of K (and vice versa): given $x \in K^\circ$ let F_x be the face of K defined by $F_x = \{y \in K \mid \langle x, y\rangle = 1\}$. More generally, if F is a face of K, we can define the corresponding exposed face F^Δ of the polar K° by $F^\Delta = \{y \in K^\circ \mid \langle x, y\rangle = 1 \quad \text{for all} \quad x \in F\}$. We observe that $(F^\Delta)^\Delta$ is equal to F if and only if F is exposed.

We can similarly define the notion of a polar cone. Let $C \subset \mathbb{R}^n$ be a convex cone. We note that if for some $x \in \mathbb{R}^n$ we have $\langle x, y\rangle \leq 1$ for all $y \in C$, then it follows that $\langle x, y\rangle \leq 0$ for all $y \in C$. Accordingly we define the *polar cone* C° as

$$C^\circ = \{x \in \mathbb{R}^n \mid \langle x, y\rangle \leq 0 \quad \text{for all} \quad y \in C\}.$$

The dual cone C^* is defined as the negative of the polar cone:

$$C^* = \{x \in \mathbb{R}^n \mid \langle x, y\rangle \geq 0 \quad \text{for all} \quad y \in C\}.$$

We note that the dual cone is often defined as a subset of the dual space $(\mathbb{R}^n)^*$:

$$C^* = \{\ell \in (\mathbb{R}^n)^* \mid \ell(y) \geq 0 \quad \text{for all} \quad y \in C\}.$$

Here we used an explicit identification of the dual space $(\mathbb{R}^n)^*$ with $\mathbb{R}^n$ via the inner product $\langle\ ,\rangle$. We can similarly state a biduality theorem for cones.

Theorem A.20. *Let C be a closed convex cone in $\mathbb{R}^n$. Then*

$$(C^\circ)^\circ = C \quad \text{and} \quad (C^*)^* = C.$$

A.3.5 Cone of Positive Semidefinite Matrices

Let $\mathcal{S}^n_+$ denote the cone of positive semidefinite $n \times n$ matrices. It is easy to show that $\mathcal{S}^n_+$ is a closed, pointed cone and it is full dimensional in $\mathcal{S}^n$. We define an inner product on $\mathcal{S}^n$ as follows: $\langle A, B\rangle = \operatorname{Tr}(AB)$. It is not hard to show that the cone $\mathcal{S}^n_+$ is self-dual.

Proposition A.21. *With the inner product $\langle A, B\rangle = \operatorname{Tr}(AB)$ for $A, B \in \mathcal{S}^n$ we have*

$$\left(\mathcal{S}^n_+\right)^* = \mathcal{S}^n_+.$$

From diagonalization of symmetric matrices we see that any positive semidefinite matrix $A \in \mathcal{S}_+^n$ can be written as a sum of rank 1 positive semidefinite matrices. Thus we see that the extreme rays of $\mathcal{S}_+^n$ are the rank 1 positive semidefinite matrices. Now let V be a linear subspace of $\mathbb{R}^n$. Let F_V be the set of all positive semidefinite matrices A such that $V \subseteq \ker A$. It is easy to show that F_V is a face of $\mathcal{S}_+^n$. It also happens that all faces of $\mathcal{S}_+^n$ have this form.

Theorem A.22. *Let F be a face of $\mathcal{S}_+^n$; then $F = F_V$ for some subspace V of $\mathbb{R}^n$.*

Using diagonalization of symmetric matrices again it follows that the face F_V is isomorphic to the cone of positive semidefinite matrices of dimension $n - \operatorname{codim} V$. Therefore faces of $\mathcal{S}_+^n$ have dimension $\binom{k+1}{2}$ for some $0 \leq k \leq n$. We note that if V and W are linear subspaces of $\mathbb{R}^n$ and $V \subseteq W$, then $F_W \subseteq F_V$. From this we obtain the following description of the face lattice $\mathcal{F}\left(\mathcal{S}_+^n\right)$.

Corollary A.23. *The face lattice $\mathcal{F}\left(\mathcal{S}_+^n\right)$ is isomorphic to the lattice of linear subspaces of $\mathbb{R}^n$ ordered by reverse inclusion.*

We also have the following important corollary.

Corollary A.24. *Let A, B be positive semidefinite matrices. Then $\langle A, B\rangle = 0$ if and only if $AB = 0$.*

Proof. Suppose that $AB = 0$. Then $\langle A, B\rangle = \operatorname{Tr}(AB) = 0$. Now suppose that $\langle A, B\rangle = 0$. We can write $B = \sum_{i=1}^k R_i$, where R_i are positive semidefinite rank 1 matrices. Then we have $\langle A, B\rangle = \sum_{i=1}^k \langle A, R_i\rangle = 0$. Since the cone $\mathcal{S}_+^n$ is self-dual, we know that $\langle A, R_i\rangle \geq 0$ and therefore $\langle A, R_i\rangle = 0$ for all i. Since matrices R_i have rank 1 we can write $R_i = v_i v_i^T$ for some vectors $v_i \in \mathbb{R}^n$. Therefore we see that $\langle A, R_i\rangle = v_i^T A v_i = 0$, and since $A \succeq 0$ we see that v_i is in the kernel of A. Therefore we see that $AR_i = Av_i v_i^T = 0$ for all i. Thus we have $AB = 0$. □

A.3.6 Dimensional Inequalities

It is often of great interest to find a *low rank* positive semidefinite matrix given some linear conditions on the entries of a matrix. While existence of a positive semidefinite matrix subject to linear constraints can be solved via semidefinite programming, the existence of a solution of low rank is a nonconvex problem and thus quite challenging. It is therefore of interest to find some theoretical guarantees on the existence of low rank solutions, given that a positive semidefinite solution exists. We state the following bounds discovered and rediscovered independently by several authors [1].

Theorem A.25. *Let A be an affine subspace of $\mathcal{S}_+^n$ such that the intersection $A \cap \mathcal{S}_+^n$ is nonempty and $\operatorname{codim} A \leq \binom{r+2}{2} - 1$ for some nonnegative integer r. Then there is a matrix $X \in \mathcal{S}_+^n \cap A$ such that $\operatorname{rank} X \leq r$.*

This bound is sharp in general, but it was improved by Barvinok in the case where the intersection $A \cap \mathcal{S}_+^n$ is bounded [1].

Theorem A.26. *Suppose that $r \geq 0$ and $n \geq r+2$. Let A be an affine subspace of $\mathcal{S}_+^n$ such that* $\operatorname{codim} A = \binom{r+2}{2}$. *Suppose that the intersection $A \cap \mathcal{S}_+^n$ is nonempty and bounded. Then there is a matrix $X \in \mathcal{S}_+^n \cap A$ such that* $\operatorname{rank} X \leq r$.

A.4 Algebra of Polynomials and Ideals

There are excellent books for the basics of commutative algebra, algebraic geometry, and real algebraic geometry used in this book. For polynomials, ideals, Gröbner bases, and basic algebraic geometry we refer the reader to [7], an introduction to these topics at the undergraduate level. For basic real algebraic geometry concepts such as semialgebraic sets and the Tarski–Seidenberg quantifier elimination, see [12]. What we provide below is a brief tour through some of the algebraic themes that arise in this book with the goal of giving the absolute newcomer a quick grasp of the concepts. For a more serious appreciation of these topics, the reader is referred to the above-mentioned books.

A.4.1 Monomials, Polynomials, and the Polynomial Ring

A *monomial* in the n variables $x_1, \ldots, x_n$ (abbreviated as x) is a product $x^a := x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$, where $a = (a_1, \ldots, a_n) \in \mathbb{N}^n$. A *polynomial* in $x_1, \ldots, x_n$ with coefficients in a *field* k is a finite linear combination of the form $f := \sum c_a x^a$, where $c_a \in k$. A monomial x^a is in the *support* of f if $c_a \neq 0$ in the expression $f = \sum c_a x^a$. The *degree* of $f = \sum c_a x^a$ is the maximum L_1-norm of the vectors a that appear as exponents of monomials in the support of f. The usual fields considered in this book are the set of real numbers denoted as $\mathbb{R}$ and the set of complex numbers denoted as $\mathbb{C}$. In what follows, we assume that the field k is either $\mathbb{C}$ or $\mathbb{R}$. The *polynomial ring* $k[x] := k[x_1, \ldots, x_n]$ is the set of all polynomials in $x_1, \ldots, x_n$ with coefficients in k. It is endowed with the two binary operations of addition and multiplication of pairs of polynomials.

Groups, rings, and fields are basic objects in abstract algebra that satisfy an increasing list of properties. See, for instance, [8] for definitions and examples. A binary operation $\star$ on a set S is

- *associative* if $(f \star g) \star h = f \star (g \star h)$ for all $f, g, h \in S$, and
- *commutative* if $f \star g = g \star f$ for all $f, g \in S$.

The pair $(S, \star)$

- has an *identity* if there exists an element $e \in S$ such that $f \star e = e \star f = f$ for all $f \in S$, and
- has *inverses* if for each $f \in S$, there exists an element $f^{-1} \in S$ such that $f \star f^{-1} = f^{-1} \star f = e$.

Definition A.27.

- *A set G with a binary operation $\star$ is a* group *if $\star$ is associative and $(G, \star)$ has an identity and inverses. If in addition, $\star$ is commutative in G, then G is called a* commutative group.

- *A set R with two binary operations $+$ (addition) and $\cdot$ (multiplication) is a* ring *if $(R,+)$ is a commutative group, and $(R,\cdot)$ is associative and $\cdot$ distributes over $+$ in the sense that $f\cdot(g+h)=f\cdot g+f\cdot h$ for all $f,g,h\in R$. If $(R,\cdot)$ has an identity and/or $(R,\cdot)$ is commutative, then R is a ring with identity and/or commutative.*
- *A* field *is a ring $(F,+,\cdot)$ in which $(F,\cdot)$ is also a commutative group with identity.*

The set of integers under addition and multiplication, $(\mathbb{Z},+,\cdot)$, forms a commutative ring with identity: $(\mathbb{Z},+)$ is a commutative group (with 0 as its additive identity and for each $z\in\mathbb{Z}$, $-z$ is the additive inverse of z), and 1 is the multiplicative identity in $(\mathbb{Z},+,\cdot)$. No element $z\in\mathbb{Z}, z\neq\pm1$ has a multiplicative inverse. On the other hand, $(\mathbb{R},+,\cdot)$ and $(\mathbb{C},+,\cdot)$ are fields. The set of $n\times n$ matrices under matrix addition and multiplication forms a noncommutative ring with identity.

The polynomial ring $k[x]$ is a commutative ring with identity under addition and multiplication of pairs of polynomials. The empty monomial $x_1^0\cdots x_n^0=1\in k[x]$ and hence k is a subset of $k[x]$ and is called the set of scalars in $k[x]$. It is customary to denote $f\cdot g$ as just fg when the multiplication operation is clear. The ring $k\langle x\rangle$ denotes the *free ring* where the variables $x_1,\ldots,x_n$ do not commute; i.e., the relation $x_ix_j=x_jx_i$ is not assumed. The free ring (and also $k[x]$) is an example of an *algebra* which is a ring that is also a vector space over its field of scalars. Hence it is often called the *free algebra* in n variables over k. This noncommutative ring plays a central role in Chapter 8.

A.4.2 Polynomial Ideals, Gröbner Bases, and Quotient Rings

Definition A.28.

1. *A subset $I\subset k[x]$ is an* ideal *if it satisfies the following properties:*
   - $0\in I$.
   - *If $f,g\in I$, then $f+g\in I$.*
   - *If $f\in I$ and $h\in k[x]$, then $hf\in I$.*
2. *The ideal generated by $f_1,\ldots,f_t\in k[x]$ is the set $I=\left\{\sum_{i=1}^t h_if_i\,:\,h_i\in k[x]\right\}$, denoted as $\langle f_1,\ldots,f_t\rangle$.*

Check that $\langle f_1,\ldots,f_t\rangle$ is an ideal in $k[x]$. A simple example of an ideal in the polynomial ring $\mathbb{R}[x_1,x_2]$ is the set of all polynomials that evaluate to 0 on the point $(0,0)$. This ideal consists of all polynomials of the form x_1f+x_2g, where $f,g\in k[x]$ and hence equals $\langle x_1,x_2\rangle$. An ideal $I\subset k[x]$ is *finitely generated* if it is generated by a finite set of polynomials in $k[x]$. A generating set of an ideal I is called a *basis* of I. An ideal can have bases of different cardinalities and, unlike a vector space basis, an ideal basis is just a generating set with no independence requirements.

Theorem A.29 (Hilbert's basis theorem). *If k is a field, then every ideal in $k[x]$ is finitely generated (has a finite basis).*

Here are two important types of ideals.

Definition A.30. *A polynomial $f = \sum c_a x^a$ is* homogeneous *if all monomials in its support have the same degree. An ideal $I \subset k[x]$ is* homogeneous *if it is generated by homogeneous polynomials.*

Definition A.31. *An ideal I is* principal *if it is generated by a single polynomial.*

Gröbner bases are special bases for a polynomial ideal. They enable many algorithms in computational algebraic geometry.

Definition A.32. *A* term order $\succ$ *on $k[x]$ is a total ordering on the monomials in $k[x]$ such that*

- $1 \prec x^a$ *for all $a \neq 0$, and*
- *if $x^a \succ x^b$ then $x^a x^c \succ x^b x^c$ for all monomials x^c.*

A common example of a term order is the *lexicographic/dictionary* order with $x_1 \succ x_2 \succ \cdots \succ x_n$ defined as $x^a \succ x^b$ if and only if the left most nonzero entry in $a - b$ is positive. Note that there are $n!$ lexicographic term orders on $k[x]$. A term order needed in Chapter 7 is the *total degree order* which first sorts monomials by degree and then breaks ties using a fixed term order such as the above lexicographic order. More precisely, $x^a \succ x^b$ if and only if either $\deg(x^a) > \deg(x^b)$ or $\deg(x^a) = \deg(x^b)$ and x^a is lexicographically larger than x^b.

Definition A.33. *Fix a term order $\succ$ on $k[x]$.*

- *The* initial term $\mathrm{in}_\succ(f)$ *of a polynomial $f = \sum c_a x^a \in k[x]$ with respect to $\succ$ is that monomial $c_a x^a$ with $c_a \neq 0$ such that $x^a \succ x^b$ for all other monomials x^b in the support of f. The monomial x^a is called the* initial monomial *of f.*
- *The* initial ideal $\mathrm{in}_\succ(I)$ *is the ideal generated by the initial monomials of all polynomials in I.*

By Hilbert's basis theorem, the initial ideal $\mathrm{in}_\succ(I)$ is finitely generated. In fact, it has a unique set of minimal generators that are all monomials.

Definition A.34. *A* Gröbner basis $G_\succ$ *of a polynomial ideal $I \subset k[x]$ with respect to the term order $\succ$ is a finite set of polynomials $g_1, \ldots, g_t \in I$ such that $\langle \mathrm{in}_\succ(g_1), \ldots, \mathrm{in}_\succ(g_t) \rangle = \mathrm{in}_\succ(I)$.*

Each term order $\succ$ gives rise to a *reduced* Gröbner basis of I which is unique. The Gröbner bases returned by a computer algebra package such as Macaulay2 are usually reduced. In the 1960s Buchberger provided an algorithm to find a Gröbner

basis of an ideal given a term order. This algorithm underlies the Gröbner basis functionality in modern computer algebra packages such as Macaulay2, SINGULAR, Maple, Mathematica, etc.

Example A.35. An example of a reduced Gröbner basis with respect to the total degree ordering was given in Chapter 7. Consider the ideal

$$I = \langle x^4 - y^2 - z^2, x^4 + x^2 + y^2 - 1 \rangle.$$

Using Macaulay2 [10] one can calculate a total degree reduced Gröbner basis of I as follows:

```
Macaulay2, version 1.3

i1 : R = QQ[x,y,z,Weights => {1,1,1}];
i2 : I = ideal(x^4-y^2-z^2, x^4+x^2+y^2-1);
i3 : G = gens gb I
o3 = | x2+2y2+z2-1 4y4+4y2z2+z4-5y2-3z2+1 |
```

which says that a total degree Gröbner basis consists of the two polynomials

$$x^2 + 2y^2 + z^2 - 1 \quad \text{and} \quad 4y^4 + 4y^2z^2 + z^4 - 5y^2 - 3z^2 + 1.$$

The reduced Gröbner basis of I would have the property that no initial term of an element is divisible by the initial term of another element and that all initial terms have unit coefficients. Hence the reduced Gröbner basis of I is

$$\left\{ x^2 + 2y^2 + z^2 - 1, \;\; y^4 + y^2z^2 + \frac{1}{4}(z^4 - 5y^2 - 3z^2 + 1) \right\}.$$

In particular, the initial ideal of I with respect to this term order is $\langle x^2, y^4 \rangle$. Check that both elements in the Gröbner basis lie in the ideal I. ∎

Gröbner bases enable a multitude of computations with ideals such as checking whether a polynomial lies in an ideal (*ideal membership*), checking whether an ideal equals the whole ring, finding all roots of a system of polynomial equations, finding the intersection of two ideals, etc. Ideal membership relies on a multivariate *division algorithm* that computes the remainder (called a *normal form*) of a polynomial f with respect to a Gröbner basis. A polynomial f lies in I if and only if the normal form of f (with respect to any reduced Gröbner basis of I) is zero. This in turn relies on the fact that the normal form of a polynomial with respect to a reduced Gröbner basis of I is unique.

Example A.36. The normal form of the monomial x^2y with respect to the Gröbner basis in Example A.35 is obtained by successively dividing out the initial monomial $\text{in}_\succ(g)$ of an element $g := \text{in}_\succ(g) - g'$ in the reduced Gröbner basis from x^2y and multiplying with g'. Let $g_1 := x^2 + 2y^2 + z^2 - 1$ and $g_2 := y^4 + y^2z^2 + \frac{1}{4}(z^4 - 5y^2 - 3z^2 + 1)$. Then x^2y can be divided by g_1 to give $-2y^3 - yz^2 + y$. The resulting initial term $-2y^3$ cannot be divided by either $in_\succ(g_1)$ or $\text{in}_\succ(g_2)$, which implies that the normal form of x^2y is $-2y^3 - yz^2 + y$. ∎

Given an ideal I in a polynomial ring $k[x]$, one can compute the *quotient ring* $k[x]/I$ which consists of all equivalence classes of polynomials in $k[x]$ mod the ideal I. Given two polynomials $f, g \in k[x]$, f is equivalent to g mod I if $f - g \in I$. This is denoted as $f \cong g \bmod I$, and the equivalence class of f mod I is denoted as $f + I$. This notion is a generalization of the familiar modular arithmetic in the ring of integers, where we say that $z, z' \in \mathbb{Z}$ are equivalent mod a fixed integer p if $z - z'$ is an integer multiple of p. In this case the ideal I (in the ring of integers $\mathbb{Z}$) is the ideal generated by p, namely the set consisting of all integer multiples of p. If f' is the normal form of a polynomial f with respect to a reduced Gröbner basis of an ideal I in $k[x]$, then $f - f' \in I$ and hence $f \cong f' \bmod I$. Since the normal form of a polynomial with respect to a reduced Gröbner basis is unique, if $f - g \in I$, then the normal form of $f - g$ is zero, which implies that both f and g have the same normal form. Hence every equivalence class of polynomials mod I can be represented by the unique normal form of all the elements in that class with respect to a fixed reduced Gröbner basis of I.

Example A.37. In Example A.35, the equivalence class of $x^2y \bmod I$ consists of all polynomials $g \in \mathbb{Q}[x, y, z]$ such that $x^2y - g \in I$. In other words, $x^2y + I$ is the set of all $g \in \mathbb{Q}[x, y, z]$ with normal form $-2y^3 - yz^2 + y$ with respect to the reduced Gröbner basis

$$\left\{ g_1 := x^2 + 2y^2 + z^2 - 1, g_2 := y^4 + y^2z^2 + \frac{1}{4}(z^4 - 5y^2 - 3z^2 + 1) \right\}. \quad \blacksquare$$

The quotient ring $k[x]/I$ is a k-vector space. Addition in the ring is defined as $(f + I) + (g + I) = (f + g) + I$ and scalar multiplication as $\alpha(f + I) = \alpha f + I$ for all $\alpha \in k$. A primary use of Gröbner bases is that they provide a vector space basis for $k[x]/I$ in the following sense. Fix a term order $\succ$ on $k[x]$ and consider the initial ideal $\text{in}_\succ(I)$ of the ideal I. Recall that this initial ideal is generated by monomials. The monomials in $k[x]$ that do not lie in $\text{in}_\succ(I)$ are called the *standard monomials* of $\text{in}_\succ(I)$. The equivalence classes $m + I$ as m varies over the standard monomials of $\text{in}_\succ(I)$ form a vector space basis of $k[x]/I$. Buchberger's algorithm for Gröbner bases was motivated by the quest to find vector space bases for $k[x]/I$. It is easy to see why the equivalence classes of standard monomials provide a vector space basis for $k[x]/I$. We saw earlier that once a term order $\succ$ is fixed, every equivalence class $f + I$ has a unique representative $f' + I$, where f' is the normal form of f with respect to the reduced Gröbner basis $G_\succ$ of I corresponding to $\succ$. Note that f' cannot be divided by $\text{in}_\succ(g)$ for any $g \in G_\succ$ and hence all its monomials are standard with respect to $\text{in}_\succ(I)$. This shows that the elements $m + I$ span $k[x]/I$. If a collection of them are linearly dependent, then there exists standard monomials $m_1, \ldots, m_t$ and scalars $\alpha_1, \ldots, \alpha_t$ such that $\sum \alpha_i(m_i + I) = 0 + I$, or equivalently, $\sum \alpha_i m_i \in I$. However, if $\sum \alpha_i m_i \in I$, then its normal form with respect to $G_\succ$ is zero which implies that some m_i is divisible by some $\text{in}_\succ(g)$ for $g \in G_\succ$, which is a contradiction.

Example A.38. The vector space $\mathbb{Q}[x, y, z]/I$ for the ideal in Example A.35 has infinite dimension. The initial ideal of the total degree order $\succ$ used in this example

is $\text{in}_\succ(I) = \langle x^2, y^4 \rangle$. Hence the standard monomials of this initial ideal are all monomials in x, y, z that are not divisible by x^2 and y^4. There are infinitely many such monomials since all powers of z are standard. Regardless, an infinite basis of $\mathbb{Q}[x, y, z]/I$ consists of $m+I$ as m varies over the standard monomials of $\text{in}_\succ(I)$. ∎

A.4.3 Algebraic Varieties

Definition A.39. *Given an ideal $I = \langle f_1, \ldots, f_t \rangle \subset k[x]$, its* affine variety *in k^n is the set $V_k(I) := \{p \in k^n : f(p) = 0 \text{ for all } f \in I\}$.*

It can be checked that $V_k(I) = \{p \in k^n : f_1(p) = \cdots = f_t(p) = 0\}$ and is hence the set of solutions (zeros) in k^n of the system of polynomial equations $f_1(x) = \cdots = f_t(x) = 0$. The affine variety of a principal ideal $\langle f \rangle \subset k[x]$ is called a *hypersurface* in k^n and denoted simply as $V_k(f)$.

Example A.40. The affine variety of the ideal $\langle x_i^2 - x_i \text{ for all } i = 1, \ldots, n \rangle \subset \mathbb{R}[x_1, \ldots, x_n]$ is the set of all 0/1 vectors in $\mathbb{R}^n$. ∎

If f is a homogeneous polynomial in $k[x]$, then for every $p \in k^n$ such that $f(p) = 0$, we also have that $f(\lambda p) = 0$ for all $\lambda \in k$. Hence, solutions of I come in lines through the origin. Hence, it makes sense to declare all points on a line through the origin in k^n as being equivalent. This leads to *projective geometry*, where we replace k^n with the *projective space* $\mathbb{P}_k^{n-1}$ whose points are in bijection with the distinct lines through the origin in k^n. The *homogeneous coordinates* of the point in $\mathbb{P}_k^{n-1}$ corresponding to the line spanned by $(x_1, \ldots, x_n)$ is denoted as $(x_1 : \cdots : x_n)$ to denote that it is unique only up to scalar multiplication. For details on the construction of projective spaces, see Chapter 8 in [7]. If a polynomial is not homogeneous, then it is not true that $p(x) = 0$ implies $p(\lambda x) = 0$ for all $\lambda \neq 0$.

Definition A.41. *The* projective variety *of a homogeneous ideal $I \subset k[x]$ is $\{p \in \mathbb{P}_k^{n-1} : f(p) = 0 \text{ for all } f \in I\}$.*

We do not introduce any notation for projective varieties here as we will not discuss them in this appendix. Chapter 8 in [7] gives an introduction to projective varieties and their relationship to affine varieties. As for affine varieties and their ideals, Gröbner bases play an important role in computations involving projective varieties and their (homogeneous) ideals.

Example A.42. The homogeneous principal ideal $I = \langle yz - x^2 \rangle \subset \mathbb{C}[x, y, z]$ contains all lines spanned by the points $(t, t^2, 1)$, $t \in \mathbb{C}$, in its affine variety in $\mathbb{C}^3$. Its projective variety is $\{(t : t^2 : 1) : t \in \mathbb{C}\} \cup \{(0 : u : 0) : u \in \mathbb{C}\} \subset \mathbb{P}_{\mathbb{C}}^2$. ∎

A field k is *algebraically closed* if every polynomial in $k[x]$ has all its roots in k^n. The field $\mathbb{C}$ is algebraically closed while $\mathbb{R}$ is not. Every ideal $I \subset k[x]$ has an affine variety $V_k(I) \subset k^n$, although different ideals can have the same affine variety. For instance, both $\langle x, y \rangle$ and $\langle x^2, y^2 \rangle$ in $\mathbb{C}[x, y]$ have the affine variety $\{(0, 0)\}$ in $\mathbb{C}^2$.

Given a variety $W \subset k^n$, its *vanishing ideal*,

$$I(W) := \{f \in k[x] \,:\, f(p) = 0 \text{ for all } p \in W\}$$

is the set of all polynomials in $k[x]$ that vanish on W. Check that $I \subseteq I(V_k(I))$ and that $V_k(I(V_k(I))) = V_k(I)$. The ideal $I(V_k(I))$ is the largest ideal with the affine variety $V_k(I)$. This vanishing ideal has the important property that if f^m belongs to it, then so does f since $f^m(p) = 0$ for all $p \in V_k(I)$ implies that $f(p) = 0$ for all $p \in V_k(I)$.

Definition A.43. *The* radical *of an ideal $I \subset k[x]$ is*

$$\sqrt{I} := \{f \in k[x] \,:\, f^m \in I \;\; \textit{for some positive integer } m\}.$$

An ideal I is radical if $I = \sqrt{I}$.

The radical $\sqrt{I}$ is an ideal and both I and $\sqrt{I}$ have the same affine variety. Further, the vanishing ideal $I(V_k(I))$ is a radical ideal. The following theorem shows that when k is an algebraically closed field, there is a bijection between radical ideals in $k[x]$ and affine varieties in k^n.

Theorem A.44 (Hilbert's strong Nullstellensatz). *If k is an algebraically closed field, then $I(V_k(I)) = \sqrt{I}$.*

The following example points out the importance of k being algebraically closed in the above Nullstellensatz.

Example A.45. The ideal $I = \langle x^2 + y^2 \rangle \subset \mathbb{C}[x, y]$ is radical. Its affine variety in $\mathbb{R}^2$ is $\{(0, 0)\}$, whose vanishing ideal is $J = \langle x, y \rangle$ and $J \neq I$. On the other hand, the affine variety of I in $\mathbb{C}^2$ consists of the two lines $x = \pm iy$ whose vanishing ideal is I. ∎

There is a strong Nullstellensatz for projective varieties as well that has the same statement. However, there is a *weak Nullstellensatz* that characterizes empty varieties whose statements are different for affine and projective varieties. We refer the reader to [7, Chapter 8] for details.

Theorem A.46 (Hilbert's weak Nullstellensatz). *Let k be an algebraically closed field.*

1. *If I is an ideal in $k[x]$, then its affine variety $V_k(I) \subseteq k^n$ is empty if and only if $I = k[x]$.*
2. *If I is a homogeneous ideal in $k[x]$, then its projective variety in $\mathbb{P}_k^{n-1}$ is empty if and only if for each $i = 1, \ldots, n$, there is a monomial $x_i^{m_i} \in I$ where m_i is some nonnegative integer.*

To end this subsection, we briefly discuss the notions of dimension, degree, and singular points of an algebraic variety. These notions are too subtle to be explained

correctly here and we refer the reader to [7, Chapter 9]. Dimension and degree of a variety can be computed from an algebraic entity called the *Hilbert polynomial* of the vanishing ideal of the variety. A key feature of Gröbner basis theory is that an ideal I and all its initial ideals have the same Hilbert polynomial and the polynomial has a combinatorial expression that can be computed from the standard monomials of any of its initial ideals. Intuitively, the dimension of an ideal is the dimension of the largest component of its affine variety. For instance we expect a hypersurface in k^n to have dimension $n-1$ since it is constrained by a single polynomial. However, when the field is not algebraically closed, this intuition can be wrong. For instance, $V_{\mathbb{R}}(x^2+y^2) = \{(0,0)\}$ is a zero-dimensional variety in $\mathbb{R}^2$ while $V_{\mathbb{C}}(x^2+y^2)$ is a one-dimensional variety in $\mathbb{C}^2$.

The *degree* of a variety is also defined from the Hilbert polynomial of the vanishing ideal. Intuitively we expect that slicing an r-dimensional variety in k^n with a generic plane of dimension $n-r$ through the origin would create finitely many intersections. The number of intersection points should be constant if the plane is generic enough and is intuitively the degree of the variety. For instance, the parabola defined by $y-x^2$ has two points of intersection with a generic line through the origin saying that its degree is two, while the cubic curve $y=x^3$ cuts out a variety of degree three.

A *nonsingular* (also called *regular* or *smooth*) point p on a variety W is a point where the *tangent space* to W at p has the same dimension as the component of W containing p and hence serves as a reasonable linear approximation to this component near p. For a polynomial $f \in k[x]$, let $\nabla(f) := (\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n})$ be the *gradient* of f and $\nabla(f)(p) \in k^n$ be the evaluation of $\nabla(f)$ at $p \in k^n$. Since the structure of a variety is unchanged by translation, we may assume without loss of generality that $p = 0$. If $I(W) = \langle f_1, \ldots, f_s \rangle$, then the *tangent space* of W at p is the null space of the matrix $J(0)$ whose rows are $\nabla(f_1)(0), \ldots, \nabla(f_s)(0)$. The matrix J whose rows are the polynomials $\nabla(f_1), \ldots, \nabla(f_s)$ is called the *Jacobian matrix* of $f_1, \ldots, f_s$. Thus the rank of $J(0)$ determines whether 0 is singular on W or not. In particular, 0 is a singular point on a hypersurface $V_k(f)$ if and only if $\nabla(f)(u) = 0$.

A.4.4 Real Algebraic Geometry

A good deal of the algebraic geometry that appears in this book is over $\mathbb{R}$, which is not an algebraically closed field. As a result, many of the theorems that apply over $\mathbb{C}$ do not work in this setting making the study of real varieties and their ideals more tricky than their complex counterparts. A good introduction to the real algebraic geometry background needed in this book is [12]. We define a few of the key concepts and results.

Definition A.47. *A set $S \subset \mathbb{R}^n$ defined as $S = \{x \in \mathbb{R}^n : f_i(x) \triangleright_i 0, i = 1, \ldots, t\}$, where, for each i, $\triangleright_i$ is one of $\geq, >, =, \neq$, and $f_i(x) \in \mathbb{R}[x]$, is called a* basic semialgebraic set. *A* basic closed semialgebraic set *is a set of the form $S = \{x \in \mathbb{R}^n : f_1(x) \geq 0, \ldots, f_t(x) \geq 0\}$.*

Every basic semialgebraic set can be expressed with polynomial inequalities of the form $f(x) \geq 0$ and a single inequality $g \neq 0$. In this book we only encounter basic semialgebraic sets in which $\triangleright$ is either $>, \geq$, or $=$. Note that every real algebraic variety is a basic closed semialgebraic set.

Definition A.48. *A finite union of basic semialgebraic sets in $\mathbb{R}^n$ is called a* semialgebraic set*, and a finite union of basic closed semialgebraic sets is a closed semialgebraic set.*

Semialgebraic sets are closed under finite unions, finite intersections, and complementation. The following theorem shows that semialgebraic sets are also closed under projections, a fact that is used several times in this book. For more details see [12] and [4].

Theorem A.49 (Tarski–Seidenberg theorem). *Let $S \subset \mathbb{R}^{k+n}$ be a semialgebraic set and $\pi : \mathbb{R}^{k+n} \to \mathbb{R}^n$ be the projection map that sends $(y, x) \mapsto x$. Then $\pi(S)$ is a semialgebraic set in $\mathbb{R}^n$.*

Recall that Σ denotes the set of sums of squares polynomials in $\mathbb{R}[x]$.

Definition A.50. *The* preorder *associated with a finite set of polynomials $f_1, \ldots, f_t \in \mathbb{R}[x]$ is the set*

$$\textbf{preorder}(f_1, \ldots, f_t) := \left\{ \sum s_\sigma f_1^{\sigma_1} \cdots f_t^{\sigma_t} : \sigma = (\sigma_1, \ldots, \sigma_t) \in \{0,1\}^t \text{ and } s_\sigma \in \Sigma \right\}.$$

All the polynomials in $\textbf{preorder}(f_1, \ldots, f_t)$ are nonnegative on the basic closed semialgebraic set $S = \{x \in \mathbb{R}^n : f_1(x) \geq 0, \ldots, f_t(x) \geq 0\}$.

Definition A.51. *The* real radical *of an ideal $I = \langle f_1, \ldots, f_t \rangle$ is the ideal*

$$\sqrt[\mathbb{R}]{I} := \{f \in \mathbb{R}[x] : -f^{2m} \in \Sigma + I \text{ for some nonnegative integer } m\}.$$

The ideal I is said to be real radical *if $I = \sqrt[\mathbb{R}]{I}$.*

We conclude with the Positivstellensatz and the real Nullstellensatz that play the analogous role of Hilbert's Nullstellensatz for semialgebraic sets and real varieties.

Theorem A.52 (Positivstellensatz). *Let $f_1, \ldots, f_t \in \mathbb{R}[x]$ and $S = \{x \in \mathbb{R}^n : f_1(x) \geq 0, \ldots, f_t(x) \geq 0\}$ and T be the preorder associated to $f_1, \ldots, f_t$. For a polynomial $f \in \mathbb{R}[x]$,*

1. *$f > 0$ on S if and only if there exists $p, q \in T$ such that $pf = 1 + q$;*
2. *$f \geq 0$ on S if and only if there exists an integer $m \geq 0$ and $p, q \in T$ such that $pf = f^{2m} + q$;*

3. $f = 0$ *on* S *if and only if there exists an integer* $m \geq 0$ *such that* $-f^{2m} \in T$*;*

4. $S = \emptyset$ *if and only if* $-1 \in T$*.*

Corollary A.53 (Real Nullstellensatz). *If* I *is an ideal in* $\mathbb{R}[x]$*, then its real radical ideal* $\sqrt[\mathbb{R}]{I}$ *is the largest ideal that vanishes on* $V_{\mathbb{R}}(I)$*.*

Recall that $\sqrt{I} \subset \mathbb{R}[x]$, the radical ideal of I in $\mathbb{R}[x]$ is the largest ideal that vanishes on the complex variety $V_{\mathbb{C}}(I)$. Therefore, since $V_{\mathbb{R}}(I) \subseteq V_{\mathbb{C}}(I)$, we have that $I \subseteq \sqrt{I} \subseteq \sqrt[\mathbb{R}]{I}$.

The Positivstellensatz also gives a simple solution to Hilbert's 17th problem, which asked whether every nonnegative polynomial in $\mathbb{R}[x]$ can be written as a sum of squares of rational functions in x. This was answered in the affirmative by Artin in 1927. The two-variable case was shown by Hilbert in 1893.

Bibliography

[1] A. Barvinok. A remark on the rank of positive semidefinite matrices subject to affine constraints. *Discrete Comput. Geom.*, 25:23–31, 2001.

[2] A. Barvinok. *A Course in Convexity*, Grad. Stud. Math. 54. American Mathematical Society, Providence, RI, 2002.

[3] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar. *Convex Analysis and Optimization*. Athena Scientific, Belmont, MA, 2003.

[4] J. Bochnak, M. Coste, and M-F. Roy. *Real Algebraic Geometry*. Springer, Berlin, 1998.

[5] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, UK, 2004.

[6] R. W. Cottle. Manifestations of the Schur complement. *Linear Algebra Appl.*, 8:189–211, 1974.

[7] D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties and Algorithms*. Springer-Verlag, New York, 1992.

[8] D. S. Dummit and R. M. Foote. *Abstract Algebra*. Prentice Hall Inc., Englewood Cliffs, NJ, 1991.

[9] G. H. Golub and C. F. Van Loan. *Matrix Computations*, 3rd edition. Johns Hopkins University Press, 1996.

[10] D. R. Grayson and M. E. Stillman. Macaulay 2, a software system for research in algebraic geometry. Available at http://www.math.uiuc.edu/Macaulay2/.

[11] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, UK, 1995.

[12] M. Marshall. *Positive Polynomials and Sums of Squares*. Math. Surveys Monogr. 146, American Mathematical Society, Providence, RI, 2008.

[13] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, New Jersey, 1970.

[14] R. Schneider. *Convex Bodies: The Brunn-Minkowski Theory*. Cambridge University Press, Cambridge, UK, 1993.

[15] G. Strang. *Introduction to Linear Algebra*, 4th edition. Wellesley Cambridge Press, 2009.