1   Introduction: what is this course about?

In this course we aim to understand the properties, both mathematical and computational, of sets defined by polynomial equations and inequalities. In particular, we want to work towards methods that will enable the solution (either exact or approximate) of optimization problems with feasible sets that are defined in terms of polynomial systems. Needless to say (is it?), many problems in estimation, control, robotics, statistics, machine learning, signal processing, etc., admit simple formulations in terms of polynomial equations and inequalities. However, these formulations can be tremendously difficult to solve, and thus our methods should try to exploit as many structural properties as possible.

The computational aspects of these sets are not fully understood at the moment. In the well-known case of polyhedra, for instance, there is a well defined relationship between the geometrical properties of the set (e.g., the number of facets, or the number of extreme points) and its algebraic representation. Furthermore, polyhedral sets are preserved by natural operations (e.g., projections). None of this will generally be true for (basic) semialgebraic sets, and this causes a very interesting interaction between their geometry and their algebraic descriptions.

2   Topics

To understand better what is going on, we will embark in a journey to learn a wide variety of methods used to approach these problems. Some of our stops along the way will include:

- Linear optimization, second order cones, semidefinite programming

- Algebra: groups, fields, rings

- Univariate polynomials

- Resultants and discriminants

- Hyperbolic polynomials

- Sum of squares

- Ideals, varieties, Groebner bases, Hilbert's Nullstellensatz

- Quantifier elimination

- Real Nullstellensatz

- And much more...

We are interested in computational methods, and want to emphasize efficiency. Throughout, applications will play an important role, both as motivation and illustration of the techniques.

3 Review: convexity

A very important notion in modern optimization is that of *convexity*. To a large extent, modulo some (important) technicalities, there is a huge gap between the theoretical and practical solvability of optimization problems for which the feasible set is convex, versus those where this property fails. Recommended presentations of convex optimization from a modern viewpoint are [BV04, BTN01, BNO03], with [Roc70] being the classical treatment of convex analysis.

Unless specified otherwise, we will work on a finite-dimensional real vector space, which we will identify with $\mathbb{R}^n$. The same will extend to the corresponding dual spaces. Often, we will implicitly use the standard Euclidean inner product, thus identifying $\mathbb{R}^n$ and its dual.

Here are some relevant definitions:

Definition 1 *A set S is* convex *if $x_1, x_2 \in S$ implies $\lambda x_1 + (1 - \lambda)x_2 \in S$ for all $0 \leq \lambda \leq 1$.*

The intersection of convex sets is always convex. Given a convex set S, a point $x \in S$ is *extreme* if for any two points x_1, x_2 in S, having $x = \lambda x_1 + (1 - \lambda)x_2$ and $\lambda \in (0, 1)$ implies that $x_1 = x_2 = x$.

Example 2 *The following are examples of convex sets:*

- *The n-dimensional* hypercube *is defined by $2n$ linear inequalities:*

$$\{x \in \mathbb{R}^n \ : \ -1 \leq x_i \leq 1, \qquad i = 1, \ldots, n\}.$$

  This convex set has 2^n extreme points, namely all those of the form $(\pm 1, \pm 1, \ldots, \pm 1)$.

- *The n-dimensional* Euclidean unit ball *is defined by the inequality $x_1^2 + \cdots + x_n^2 \leq 1$. This set has an infinite number of extreme points, namely all those on the hypersurface $x_1^2 + \cdots + x_n^2 = 1$.*

- *The n-dimensional* crosspolytope *has $2n$ extreme points, namely all those whose coordinates are permutations of $(\pm 1, 0, \ldots, 0)$. It can be defined using 2^n linear inequalities, of the form*

$$\pm x_1 \pm x_2 \pm \cdots \pm x_n \leq 1.$$

All these examples actually correspond to unit balls of different norms (ℓ_∞, ℓ_2, and ℓ_1, respectively). It is easy to show that the unit ball of every norm is always a convex set. Conversely, given any full-dimensional convex set symmetric with respect to the origin, one can define a norm via the corresponding gauge *(or* Minkowski*) functional.*

One of the most important results about convex sets is the *separating hyperplane* theorem.

Theorem 3 *Given two disjoint convex sets S_1, S_2 in $\mathbb{R}^n$, there exists a nontrivial linear functional c and a scalar d such that*

$$\langle c, x \rangle \geq d \qquad \forall x \in S_1$$
$$\langle c, x \rangle \leq d \qquad \forall x \in S_2.$$

Under certain additional conditions, *strict* separation can be guaranteed. One of the most useful cases is when one of the sets is compact and the other one is closed.

Convex cones. An important class of convex sets are those that are invariant under nonnegative scalings.

Definition 4 *A set $S \subseteq \mathbb{R}^n$ is a* cone *if $\lambda \geq 0$, $x \in S \Rightarrow \lambda x \in S$.*

Definition 5 *The* dual *of a set S is $S^* := \{y \in \mathbb{R}^n : \langle y, x \rangle \geq 0 \quad \forall x \in S\}$.*

Given any set S, its dual S^* is always a closed convex cone. Duality reverses inclusion, that is, $S_1 \subseteq S_2$ implies $S_1^* \supseteq S_2^*$. If S is a closed convex cone, then $S^{**} = S$. Otherwise, S^{**} is the closure of the smallest convex cone that contains S.

A cone $\mathcal{K}$ is *pointed* if $\mathcal{K} \cap (-\mathcal{K}) = \{0\}$, and *solid* if the interior of $\mathcal{K}$ is not empty. A cone that is convex, closed, pointed and solid is called a *proper* cone. The dual set of a proper cone is also a proper cone, called the *dual cone*. An element x is in the interior of the cone K if and only if $\langle x, y \rangle > 0$, $\forall y \in K^*, y \neq 0$.

Example 6 *The* nonnegative orthant *is defined as $\mathbb{R}_+^n := \{x \in \mathbb{R}^n : x_i \geq 0\}$, and is a proper cone. The nonnegative orthant is self-dual, i.e., we have $(\mathbb{R}_+^n)^* = \mathbb{R}_+^n$.*

A proper cone $\mathcal{K}$ induces a *partial order*[1] on the vector space, via $x \preceq y$ if and only if $y - x \in \mathcal{K}$. We also use $x \prec y$ if $y - x$ is in the interior of $\mathcal{K}$. Important examples of proper cones are the nonnegative orthant, the Lorentz cone, the set of symmetric positive semidefinite matrices, and the set of nonnegative polynomials. We will discuss some of these in more detail later in the lectures and the exercises.

Example 7 *Consider the* second-order cone*, defined by $\{(x_0, x_1, \ldots, x_n) \in \mathbb{R}^{n+1} : \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}} \leq x_0\}$. This is a self-dual proper cone, and is also known as the* ice-cream*, or* Lorentz *cone.*

An interesting physical interpretation of the partial order induced by this cone appears in the theory of special relativity. In this case, the cone can be expressed (after an inconsequential rescaling and reordering) as

$$\{(x, y, z, t) \in \mathbb{R}^4 : x^2 + y^2 + z^2 \leq c^2 t^2, \quad t \geq 0\},$$

where c is a given constant (speed of light). In this case, the vector space is interpreted as the Minkowski spacetime. Given a fixed point x_0, those points x for which $x \succeq x_0$ correspond to the absolute future*, while those for which $x \preceq x_0$ are in the* absolute past*. There are, however, many points that are neither in the absolute future nor in the absolute past (for these, the causal order will depend on the observer).*

Remark 8 *Convexity has two natural definitions. The first one is the one given above, that emphasizes the "internal" aspect, in terms of convex combinations of elements of the set. Alternatively, one can look at the "external" aspect, and define a convex set as the intersection of a (possibly infinite) collection of half-spaces. The possibility of these "dual" descriptions is what enables many of the useful and intriguing properties of convex sets. In the context of convex functions, for instance, these ideas are made concrete through the use of the Legendre-Fenchel transformation.*

---

[1] A *partial order* is a binary relation $\preceq$ that is reflexive, antisymmetric, and transitive.

4 Review: linear programming

Linear programming (LP) is the problem of minimizing a linear function, subject to finitely many linear inequality constraints. An LP in standard form is written as

$$\min c^T x \qquad \text{s.t.} \quad \begin{cases} Ax & = & b \\ x & \geq & 0. \end{cases} \tag{P}$$

Every LP problem has a corresponding *dual* problem, which in this case is

$$\max b^T y \qquad \text{s.t.} \quad c - A^T y \geq 0. \tag{D}$$

Linear programming has many important properties. Among them, we mention the following ones:

Geometry of the feasible set: The feasible sets of linear programs are *polyhedra*. The geometry of polyhedra is quite well understood. In particular, the Minkowski-Weyl theorem (e.g. [BT97], or [Zie95, Section 1.1]) states that every polyhedron P is finitely generated, i.e., it can be written as
$$P = \text{conv}(u_1, \ldots, u_r) + \text{cone}(v_1, \ldots, v_s),$$
where u_i, v_i are the *vertices* and *extreme rays* of P, and the *convex hull* and *conical hull* are given by

$$\text{conv}(u_1, \ldots, u_r) = \left\{ \sum_{i=1}^{r} \lambda_i u_i \;\middle|\; \sum_{i=1}^{r} \lambda_i = 1, \quad \lambda_i \geq 0, \quad i = 1, \ldots, r \right\},$$

and

$$\text{cone}(v_1, \ldots, v_s) = \left\{ \sum_{i=1}^{s} \lambda_i v_i \;\middle|\; \lambda_i \geq 0, \quad i = 1, \ldots, s \right\}.$$

Rational solutions: Unless the problem is unbounded, the optimal solution of a linear programming problem is always achieved at extreme points of the feasible set. Since these correspond to vertices of a polyhedron, the solution can be characterized in terms of a system of linear equations, corresponding to the equations and inequalities that are active at the optimal point. Thus, if the problem description (i.e., the matrices A, b, c) is given by rational numbers, there are always extreme points that are rational and achieve the optimal cost.

Weak duality: For *any* feasible solutions x, y of (P) and (D), respectively, it always holds that:

$$c^T x - b^T y = x^T c - (Ax)^T y = x^T (c - A^T y) \geq 0, \tag{1}$$

where the last inequality follows from the feasibility conditions $x \geq 0$ and $A^T y \leq c$. Thus, from any feasible dual solution one can obtain a lower bound on the value of the primal. Conversely, primal feasible solutions give upper bounds on the value of the dual.

Strong duality: If both primal and dual problems are feasible, then they achieve exactly the same optimal value, and there exist optimal feasible solutions $x^\star, y^\star$ such that $c^T x^\star = b^T y^\star$. This is a consequence of the separation theorems for convex sets.

Complementary slackness: Strong duality, combined with (1), implies that at optimality we must have

$$x_i^\star (c - A^T y^\star)_i = 0 \qquad i = 1, \dots, n.$$

In other words, there is a correspondence between primal variables and dual inequalities, that says that whenever a primal variable is nonzero, the corresponding dual inequality must be tight. Conversely, if at optimality a dual inequality is not tight, the corresponding primal variable must vanish.

Some of these properties (which ones?) will break down as soon as we leave LP and go the more general case of conic or semidefinite programming. These will cause some difficulties, although with the right assumptions, the resulting theory will closely parallel the LP case.

Remark 9 *Among others, the software codes* `cdd` *(Komei Fukuda,* `https://www.inf.ethz.ch/personal/fukudak/cdd_home/`*),* `lrs` *(David Avis,* `http://cgm.cs.mcgill.ca/~avis/C/lrs.html`*), or* `Normaliz` *(`https://www.normaliz.uni-osnabrueck.de`) are very useful for polyhedral computations. In particular, all of them allow to convert an inequality representation of a polyhedron (usually called an H-representation) into extreme points/rays (V-representation), and viceversa. There are also MATLAB or Julia toolboxes for polyhedral computations, like* `MPT3` *(`https://www.mpt3.org`) or* `Polyhedra.jl` *(`https://github.com/JuliaPolyhedra/Polyhedra.jl`)*

References

[BNO03] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar. *Convex analysis and optimization.* Athena Scientific, Belmont, MA, 2003.

[BT97] D. Bertsimas and J. N. Tsitsiklis. *Introduction to Linear Optimization.* Athena Scientific, 1997.

[BTN01] A. Ben-Tal and A. Nemirovski. *Lectures on modern convex optimization.* MPS/SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001.

[BV04] S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004.

[Roc70] R. T. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, New Jersey, 1970.

[Zie95] G. M. Ziegler. *Lectures on polytopes*, volume 152 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1995.