Sums of Squares

Sanjay Lall
Stanford University

EE364b
April 19, 2011

polynomial programming

A familiar problem

$$
\begin{array}{lll}
\text{minimize} & f_0(x) & \\
\text{subject to} & f_i(x) \leq 0 & \text{for all } i = 1, \dots, m \\
& h_i(x) = 0 & \text{for all } i = 1, \dots, p
\end{array}
$$

in this section, objective, inequality and equality constraint functions are all *polynomials*

polynomial nonnegativity

does there exist $x \in \mathbb{R}^n$ such that $f(x) < 0$

- if not, f is called *positive semidefinite* or *PSD*

$$f(x) \geq 0 \quad \text{for all } x \in \mathbb{R}^n$$

- the problem is *NP-hard*, but decidable

certificates

does there exist $x \in \mathbb{R}^n$ such that $f(x) < 0$

- answer yes is easy to verify; exhibit x such that $f(x) < 0$
- answer no is hard; we need a *certificate* or a *witness*
  i.e, a proof that there is no feasible point

Sum of Squares Decomposition

f is nonnegative if there are polynomials $g_1, \ldots, g_s$ such that

$$f = \sum_{i=1}^{s} g_i^2$$

a checkable certificate, called a *sum-of-squares (SOS)* decomposition

- how do we find the g_i
- when does such a certificate exist?

example

we can write any polynomial as a *quadratic function of monomials*

$$
\begin{aligned}
f &= 4x^4 + 4x^3y - 7x^2y^2 - 2xy^3 + 10y^4 \\
&= \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix}^T \begin{bmatrix} 4 & 2 & -\lambda \\ 2 & -7+2\lambda & -1 \\ -\lambda & -1 & 10 \end{bmatrix} \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix} \\
&= z^T Q(\lambda) z
\end{aligned}
$$

- above equation holds for all $\lambda \in \mathbb{R}$
- if for some λ we have $Q(\lambda) \succeq 0$, then we can factorize $Q(\lambda)$

example, continued

e.g., with $\lambda = 6$, we have

$$Q(\lambda) = \begin{bmatrix} 4 & 2 & -6 \\ 2 & 5 & -1 \\ -6 & -1 & 10 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 2 & 1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} 0 & 2 & 1 \\ 2 & 1 & -3 \end{bmatrix}$$

so we have an SOS decomposition

$$f = \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix}^T \begin{bmatrix} 0 & 2 \\ 2 & 1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} 0 & 2 & 1 \\ 2 & 1 & -3 \end{bmatrix} \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix}$$

$$= \left\| \begin{bmatrix} 2xy + y^2 \\ 2x^2 + xy - 3y^2 \end{bmatrix} \right\|^2$$

$$= \left(2xy + y^2\right)^2 + \left(2x^2 + xy - 3y^2\right)^2$$

sum of squares and semidefinite programming

suppose $f \in \mathbb{R}[x_1, \ldots, x_n]$, of degree $2d$

let z be a vector of all monomials of degree less than or equal to d

f is SOS if and only if there exists Q such that

$$\boxed{\begin{aligned} Q &\succeq 0 \\ f &= z^T Q z \end{aligned}}$$

- this is an SDP in standard primal form
- the number of components of z is $\binom{n+d}{d}$
- comparing terms gives affine constraints on the elements of Q

sum of squares and semidefinite programming

if Q is a feasible point of the SDP, then to construct the SOS representation

factorize $Q = VV^T$, and write $V = \begin{bmatrix} v_1 & \dots v_r \end{bmatrix}$, so that

$$\begin{aligned} f &= z^T V V^T z \\ &= \|V^T z\|^2 \\ &= \sum_{i=1}^{r} (v_i^T z)^2 \end{aligned}$$

- one can factorize using e.g., Cholesky or eigenvalue decomposition
- the number of squares r equals the rank of Q

example

$$f = 2x^4 + 2x^3y - x^2y^2 + 5y^4$$

$$= \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix}^T \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{13} & q_{23} & q_{33} \end{bmatrix} \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix}$$

$$= q_{11}x^4 + 2q_{12}x^3y + (q_{22} + 2q_{13})x^2y^2 + 2q_{23}xy^3 + q_{33}y^4$$

so f is SOS if and only if there exists Q satisfying the SDP

$$Q \succeq 0 \qquad q_{11} = 2 \qquad 2q_{12} = 2$$
$$2q_{12} + q_{22} = -1 \qquad 2q_{23} = 0$$
$$q_{33} = 5$$

convexity

the sets of PSD and SOS polynomials are a *convex cones*; i.e.,

$$f, g \text{ PSD} \qquad \Longrightarrow \qquad \lambda f + \mu g \text{ is PSD for all } \lambda, \mu \geq 0$$

let $P_{n,d}$ be the set of PSD polynomials of degree $\leq d$
let $\Sigma_{n,d}$ be the set of SOS polynomials of degree $\leq d$

- both $P_{n,d}$ and $\Sigma_{n,d}$ are *convex cones* in $\mathbb{R}^N$ where $N = \binom{n+d}{d}$
- we know $\Sigma_{n,d} \subset P_{n,d}$, and testing if $f \in P_{n,d}$ is NP-hard
- but testing if $f \in \Sigma_{n,d}$ is an SDP (but a large one)

polynomials in one variable

if $f \in \mathbb{R}[x]$, then f is SOS if and only if f is PSD

example

all real roots must have even multiplicity, and highest coeff. is positive

$$\begin{aligned} f &= x^6 - 10x^5 + 51x^4 - 166x^3 + 342x^2 - 400x + 200 \\ &= (x-2)^2\big(x-(2+i)\big)\big(x-(2-i)\big)\big(x-(1+3i)\big)\big(x-(1-3i)\big) \end{aligned}$$

now reorder complex conjugate roots

$$\begin{aligned} &= (x-2)^2\big(x-(2+i)\big)\big(x-(1+3i)\big)\big(x-(2-i)\big)\big(x-(1-3i)\big) \\ &= (x-2)^2\big((x^2-3x-1)-i(4x-7)\big)\big((x^2-3x-1)+i(4x-7)\big) \\ &= (x-2)^2\big((x^2-3x-1)^2+(4x-7)^2\big) \end{aligned}$$

so every PSD scalar polynomial is the sum of *one or two* squares

quadratic polynomials

a quadratic polynomial in n variables is PSD if and only if it is SOS

because it is PSD if and only if

$$f = x^T Q x$$

where $Q \geq 0$

and it is SOS if and only if

$$\begin{aligned} f &= \sum_i (v_i^T x)^2 \\ &= x^T \Big(\sum_i v_i v_i^T \Big) x \end{aligned}$$

some background

In 1888, Hilbert showed that PSD=SOS if and only if

- $d = 2$, i.e., quadratic polynomials
- $n = 1$, i.e., univariate polynomials
- $d = 4, n = 2$, i.e., quartic polynomials in two variables

$n \backslash d$	2	4	6	8
1	yes	yes	yes	yes
2	yes	*yes*	no	no
3	yes	no	no	no
4	yes	no	no	no

- in general f is PSD does not imply f is SOS

some background

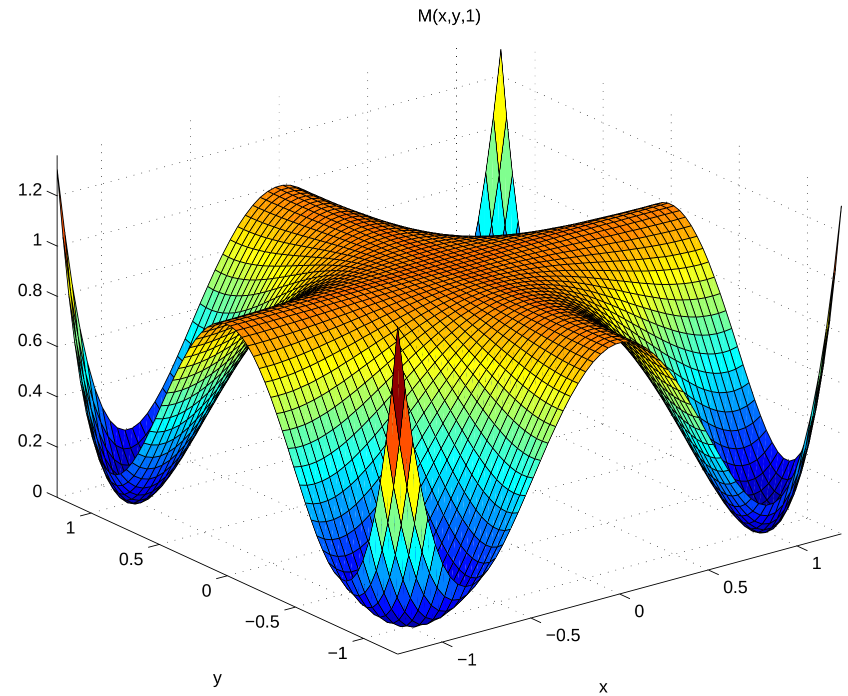
- Connections with Hilbert's 17th problem, solved by Artin: every PSD polynomial is a SOS of *rational functions*.
- If f is not SOS, then can try with gf, for some g.
  - For fixed f, can optimize over g too
  - Otherwise, can use a "universal" construction of Pólya-Reznick.

More about this later.

The Motzkin Polynomial

A positive semidefinite polynomial, that is *not* a sum of squares.

$$M(x,y) = x^2y^4 + x^4y^2 + 1 - 3x^2y^2$$



- Nonnegativity follows from the arithmetic-geometric inequality applied to $(x^2y^4, x^4y^2, 1)$
- Introduce a nonnegative factor $x^2 + y^2 + 1$
- Solving the SDPs we obtain the decomposition:

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

The Univariate Case:

$$
\begin{aligned}
f(x) &= a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_{2d} x^{2d} \\
&= \begin{bmatrix} 1 \\ x \\ \vdots \\ x^d \end{bmatrix}^T \begin{bmatrix} q_{00} & q_{01} & \cdots & q_{0d} \\ q_{01} & q_{11} & \cdots & q_{1d} \\ \vdots & \vdots & \ddots & \vdots \\ q_{0d} & q_{1d} & \cdots & q_{dd} \end{bmatrix} \begin{bmatrix} 1 \\ x \\ \vdots \\ x^d \end{bmatrix} \\
&= \sum_{i=0}^{d} \Bigg(\sum_{j+k=i} q_{jk} \Bigg) x^i
\end{aligned}
$$

- In the univariate case, the SOS condition is exactly equivalent to non-negativity.
- The matrices A_i in the SDP have a Hankel structure. This can be exploited for efficient computation.

About SOS/SDP

- The resulting SDP problem is polynomially sized (in n, for fixed d).
- By properly choosing the monomials, we can exploit structure (sparsity, symmetries, ideal structure).
- An important feature: the problem is still a SDP *if the coefficients of F are variable*, and the dependence is affine.
- Can optimize over SOS polynomials in affinely described families.

  For instance, if we have $p(x) = p_0(x) + \alpha p_1(x) + \beta p_2(x)$, we can "easily" find values of α, β for which $p(x)$ is SOS.

Global Optimization

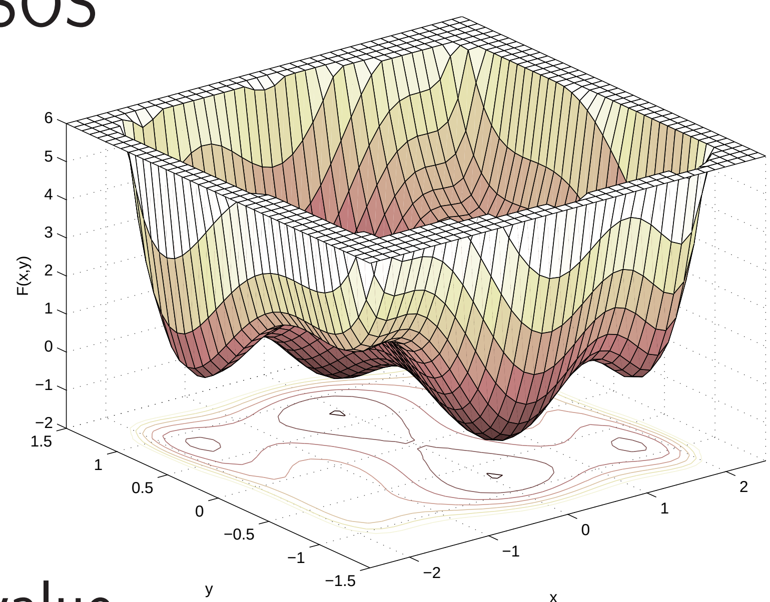
Consider the problem

$$\min_{x,y} f(x,y)$$

with

$$f(x,y) := 4x^2 - \frac{21}{10}x^4 + \frac{1}{3}x^6 + xy - 4y^2 + 4y^4$$

- Not convex. Many local minima. NP-hard.
- Find the largest γ s.t. $f(x,y) - \gamma$ is SOS
- Essentially due to Shor (1987).
- A semidefinite program (convex!).
- If exact, can recover optimal solution.
- *Surprisingly* effective.



Solving, the maximum γ is -1.0316. Exact value.

Lyapunov Example

A jet engine model

$$\begin{aligned} \dot{x} &= -y - \frac{3}{2}x^2 - \frac{1}{2}x^3 \\ \dot{y} &= 3x - y \end{aligned}$$

Try a generic 4th order polynomial Lyapunov function.

$$V(x, y) = \sum_{0 \le j+k \le 4} c_{jk} x^j y^k$$

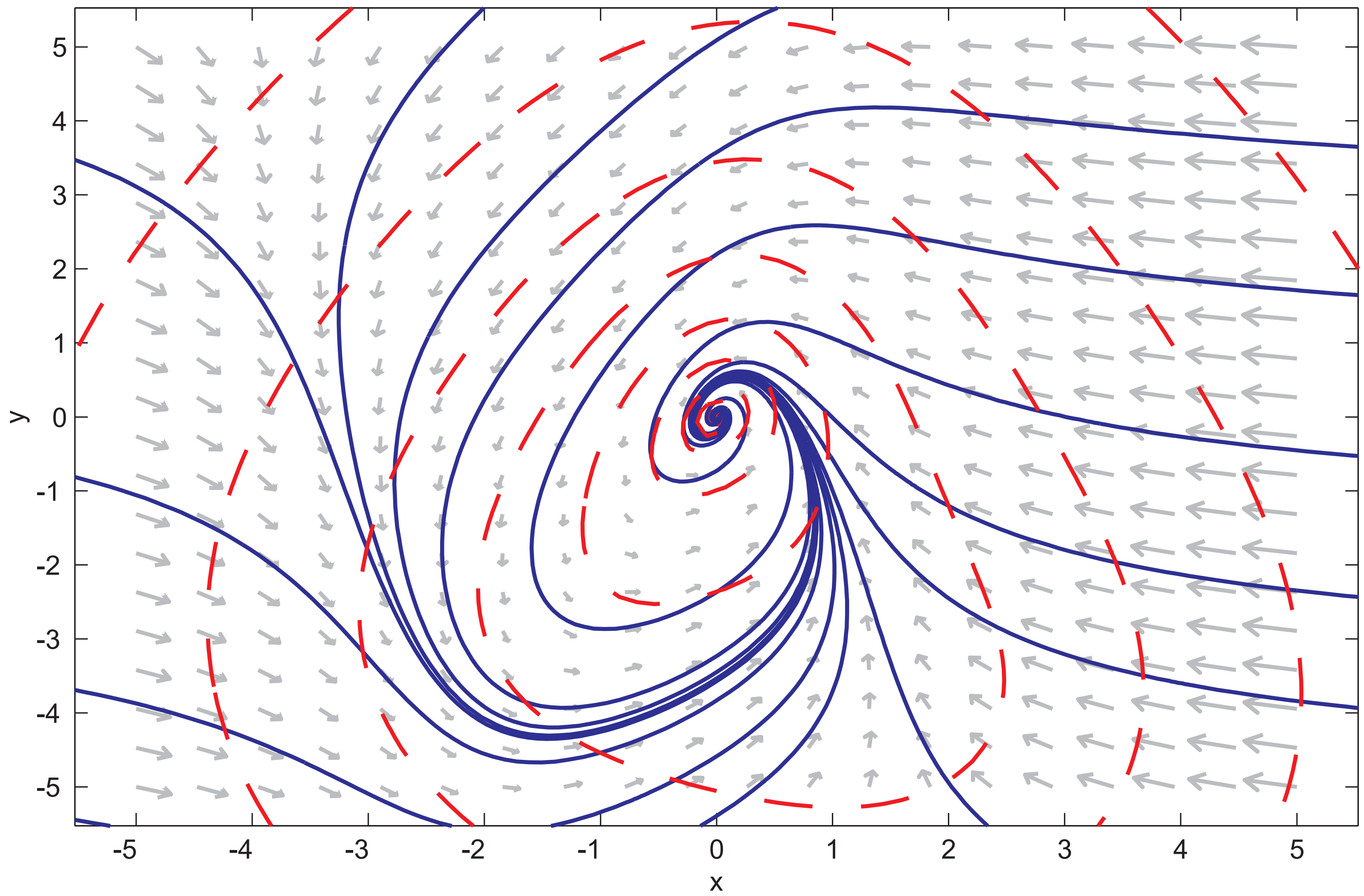
Find a $V(x, y)$ that satisfies the conditions:

- $V(x, y)$ is SOS.
- $-\dot{V}(x, y)$ is SOS.

Both conditions are affine in the c_{jk}. Can do this directly using SOS/SDP!

Lyapunov Example

After solving the SDPs, we obtain a Lyapunov function.



$$V = 4.5819x^2 - 1.5786xy + 1.7834y^2 - 0.12739x^3 + 2.5189x^2y - 0.34069xy^2$$
$$+0.61188y^3 + 0.47537x^4 - 0.052424x^3y + 0.44289x^2y^2 + 0.0000018868xy^3 + 0.090723y^4$$

Extensions

- Other linear differential inequalities (e.g. Hamilton-Jacobi).
- Many possible variations: nonlinear optimal control, parameter dependent Lyapunov functions, etc.
- Can also do local results (for instance, on compact domains).
- Polynomial and rational vector fields, or functions with an underlying algebraic structure.
- Natural extension of the SDPs for the linear case.

Automated Inference and Algebra

Automated inference is a well-known approach for formal proof systems.

Suppose $f_1(x) \geq 0$ and $f_2(x) \geq 0$, then $h(x) \geq 0$ if any of the following hold:

(i) $h(x) = f_1(x) + f_2(x)$

(ii) $h(x) = f_1(x)f_2(x)$

(iii) For any f, the function $h(x) = f(x)^2$

- We can use *algebra* to generate such *valid inequalities*
- Closure under these inference rules gives the cone of polynomials *generated* by the f_i, written $\mathbf{cone}\{f_1, f_2, \ldots, f_m\}$

The Sum-of-Squares Cone

A polynomial $f \in \mathbb{R}[x_1, \ldots, x_n]$ is called a *sum-of-squares* (SOS) if

$$f(x) = \sum_{i=1}^{r} s_i(x)^2$$

for some polynomials $s_1, \ldots, s_r$ and some $r \geq 0$

- Denote by Σ the set of SOS polynomials
- Σ is the *smallest* cone.
- This cone can be computationally characterized using semidefinite programming.
- The SOS decomposition is a simple certificate of nonnegativity of f.

The Cone

We can explicitly parameterize the cone generated by the f_i.

For example, $h \in \mathbf{cone}\{f_1, f_2, f_3\}$ if and only if

$$\boxed{h = s_1 g_1 + \cdots + s_r g_r}$$

where

$$s_i \in \Sigma \quad \text{and} \quad g_i \in \left\{ 1,\ f_1,\ f_2,\ f_3,\ f_1 f_2,\ f_2 f_3,\ f_3 f_1,\ f_1 f_2 f_3 \right\}$$

In general, every h is a linear combination of *squarefree products* of the f_i, with *SOS coefficients*

An Algebraic Dual Problem

Suppose $f_1, \ldots, f_m$ are polynomials. The primal feasibility problem is

does there exist $x \in \mathbb{R}^n$ such that
$f_i(x) \geq 0 \qquad$ for all $i = 1, \ldots, m$

The *dual feasibility problem* is

Is it true that $-1 \in \mathbf{cone}\{f_1, \ldots, f_m\}$

If the dual problem is feasible, then the primal problem is infeasible.

In fact, a result called the *Positivstellensatz* (Stengle 1974) implies the *converse*; i.e., this is a *strong duality* result.

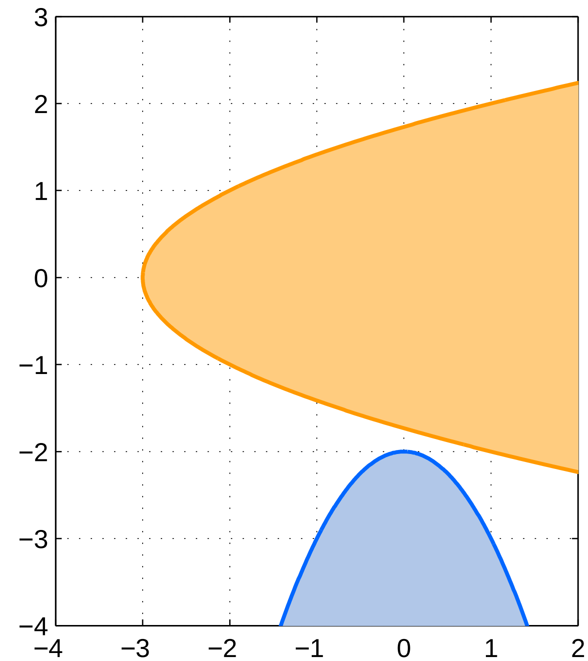
Example

Consider the feasibility problem

$$S = \{ (x,y) \in \mathbb{R}^2 \mid f(x,y) \geq 0, g(x,y) \geq 0 \}$$

where

$$f = x - y^2 + 3 \qquad g = -y - x^2 - 2$$



By the P-satz, the primal is infeasible if and only if there exist polynomials $s_0, s_1, s_2, s_3 \in \Sigma$ such that

$$-1 = s_0 + s_1 f + s_2 g + s_3 f g$$

A certificate is given by

$$s_0 = \tfrac{1}{3} + 2\big(y + \tfrac{3}{2}\big)^2 + 6\big(x - \tfrac{1}{6}\big)^2, \quad s_1 = 2, \quad s_2 = 6, \quad s_3 = 0$$

Suppose we have SOS polynomials $s_0, \ldots, s_3$ such that

$$\boxed{-1 = s_0 + s_1 f_1 + s_2 f_2 + s_3 f_1 f_2}$$

Then this is a *certificate* that there is no $x \in \mathbb{R}^n$ such that

$$f_1(x) \geq 0 \qquad \text{and} \qquad f_2(x) \geq 0$$

Positivstellensatz

The polynomials s_i give a *certificate of infeasibility* of the primal problem.

Given them, one may immediately computationally *verify* that

$$-1 = s_1 g_1 + \cdots + s_r g_r$$

and this is a *proof* of infeasibility

Finding Refutations

- Geometrically, $\mathbf{cone}\{f_1, \ldots, f_m\}$ is a *convex cone*, so testing if it contains -1 is a *convex program*.
- There is a *correspondence* between the geometric object (the *feasible set*) and the algebraic object (the *cone*).