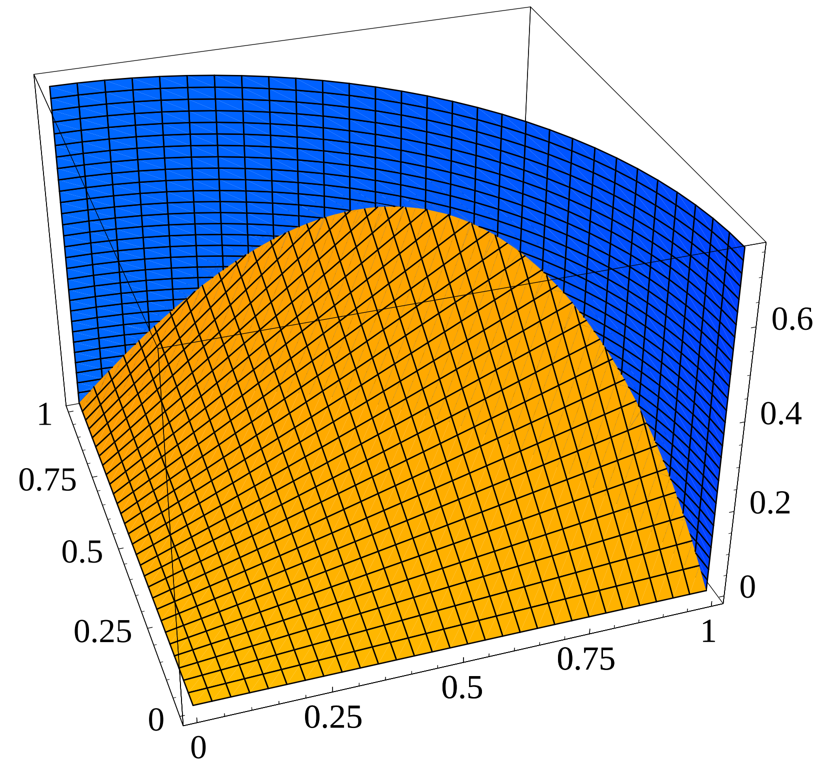
4. Algebra and Duality

- Example: non-convex polynomial optimization
- Weak duality and duality gap
- The dual is not intrinsic
- The cone of valid inequalities
- Algebraic geometry
- The cone generated by a set of polynomials
- An algebraic approach to duality
- Example: feasibility
- Searching the cone
- Interpretation as formal proof
- Example: linear inequalities and Farkas lemma

Example



$$
\begin{array}{ll}
\text{minimize} & x_1 x_2 \\
\text{subject to} & x_1 \geq 0 \\
& x_2 \geq 0 \\
& x_1^2 + x_2^2 \leq 1
\end{array}
$$

- The objective is not convex.
- The Lagrange dual function is

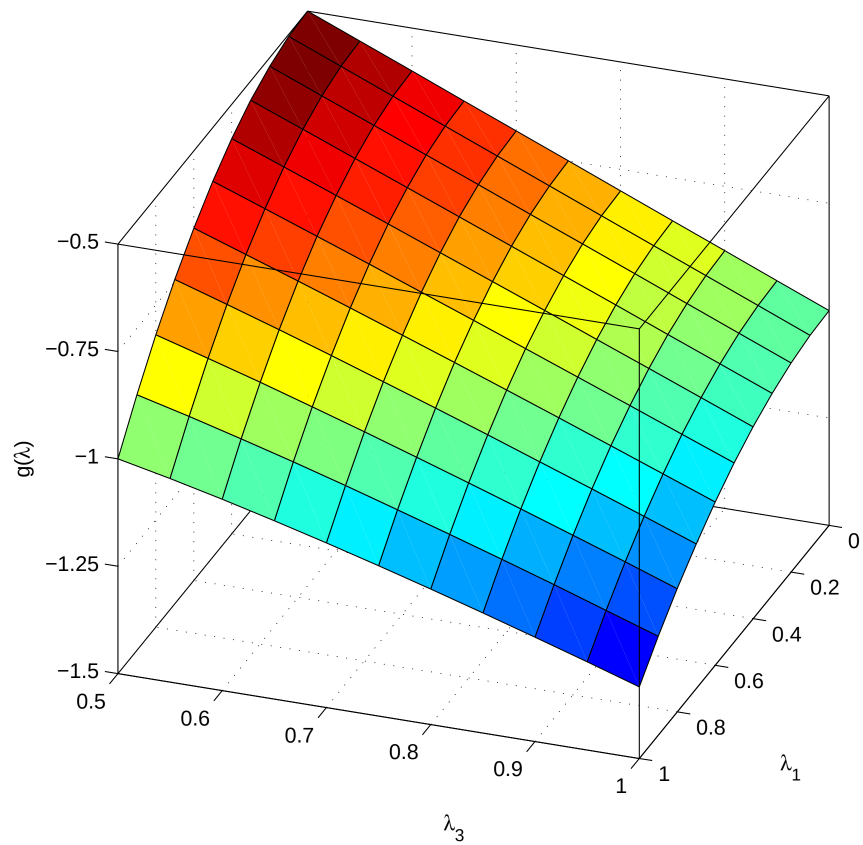
$$
g(\lambda) = \inf_x \Big(x_1 x_2 - \lambda_1 x_1 - \lambda_2 x_2 + \lambda_3 (x_1^2 + x_2^2 - 1) \Big)
$$

$$
= \begin{cases} -\lambda_3 - \frac{1}{2} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}^T \begin{bmatrix} 2\lambda_3 & 1 \\ 1 & 2\lambda_3 \end{bmatrix}^{-1} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} & \text{if } \lambda_3 > \frac{1}{2} \\ -\infty & \text{otherwise, except bdry} \end{cases}
$$

Example, continued

The dual problem is

$$\begin{array}{ll} \text{maximize} & g(\lambda) \\ \text{subject to} & \lambda_1 \geq 0 \\ & \lambda_2 \geq 0 \\ & \lambda_3 \geq \frac{1}{2} \end{array}$$



- By symmetry, if the maximum $g(\lambda)$ is attained, then $\lambda_1 = \lambda_2$ at optimality
- The optimal $g(\lambda^\star) = -\frac{1}{2}$ at $\lambda^\star = (0, 0, \frac{1}{2})$
- Here we see an example of a *duality gap*; the primal optimal is strictly greater than the dual optimal

Example, continued

It turns out that, using the Schur complement, the dual problem may be written as

$$\begin{array}{ll} \text{maximize} & \gamma \\ \text{subject to} & \begin{bmatrix} -2\gamma - 2\lambda_3 & \lambda_1 & \lambda_2 \\ \lambda_1 & 2\lambda_3 & 1 \\ \lambda_2 & 1 & 2\lambda_3 \end{bmatrix} > 0 \\ & \lambda_1 > 0 \\ & \lambda_2 > 0 \end{array}$$

In this workshop we'll see a systematic way to convert a dual problem to an SDP, whenever the objective and constraint functions are polynomials.

The Dual is Not Intrinsic

- The dual problem, and its corresponding optimal value, are *not properties of the primal feasible set and objective function alone*.
- Instead, they depend on the particular equations and inequalities used

To construct equivalent primal optimization problems with different duals:

- replace the objective $f_0(x)$ by $h(f_0(x))$ where h is increasing
- introduce new variables and associated constraints, e.g.

$$\text{minimize} \quad (x_1 - x_2)^2 + (x_2 - x_3)^2$$

is replaced by

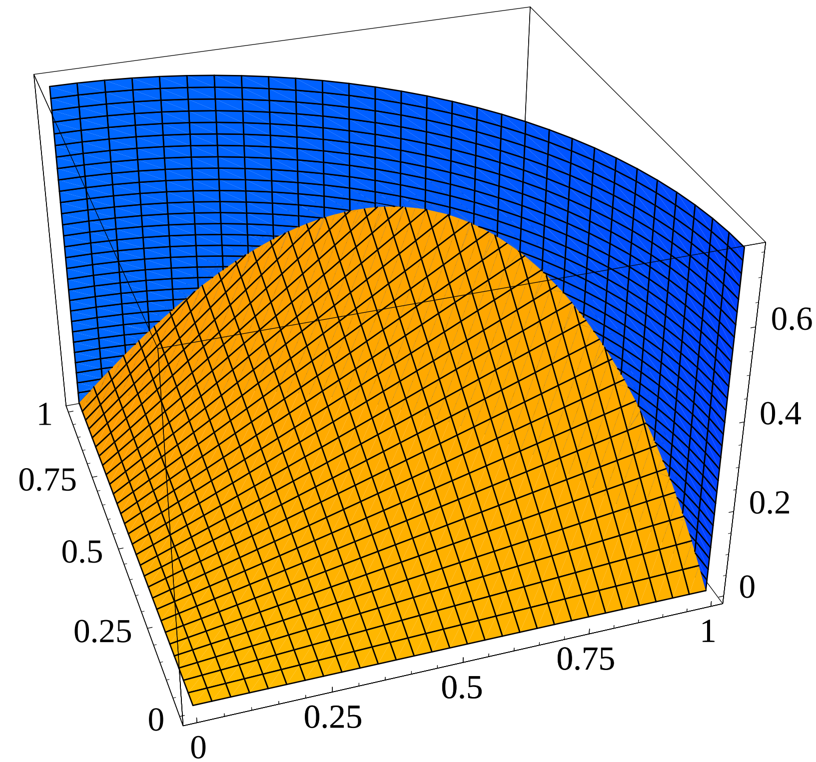
$$\begin{array}{ll} \text{minimize} & (x_1 - x_2)^2 + (x_4 - x_3)^2 \\ \text{subject to} & x_2 = x_4 \end{array}$$

- add redundant constraints

Example

Adding the redundant constraint $x_1 x_2 \geq 0$ to the previous example gives

$$\begin{array}{ll} \text{minimize} & x_1 x_2 \\ \text{subject to} & x_1 \geq 0 \\ & x_2 \geq 0 \\ & x_1^2 + x_2^2 \leq 1 \\ & x_1 x_2 \geq 0 \end{array}$$



Clearly, this has the same primal feasible set and same optimal value as before

Example Continued

The Lagrange dual function is

$$g(\lambda) = \inf_x \Big(x_1 x_2 - \lambda_1 x_1 - \lambda_2 x_2 + \lambda_3 (x_1^2 + x_2^2 - 1) - \lambda_4 x_1 x_2 \Big)$$

$$= \begin{cases} -\lambda_3 - \frac{1}{2} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}^T \begin{bmatrix} 2\lambda_3 & 1-\lambda_4 \\ 1-\lambda_4 & 2\lambda_3 \end{bmatrix}^{-1} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} & \text{if } 2\lambda_3 > 1 - \lambda_4 \\ -\infty & \text{otherwise, except bdry} \end{cases}$$

- Again, this problem may also be written as an SDP. The optimal value is $g(\lambda^\star) = 0$ at $\lambda^\star = (0, 0, 0, 1)$
- Adding redundant constraints makes the dual bound *tighter*
- This always happens! Such redundant constraints are called *valid inequalities*.

Constructing Valid Inequalities

The function $f : \mathbb{R}^n \to \mathbb{R}$ is called a *valid inequality* if

$$f(x) \geq 0 \qquad \text{for all feasible } x$$

Given a set of inequality constraints, we can generate others as follows.

(i) If f_1 and f_2 define valid inequalities, then so does $h(x) = f_1(x) + f_2(x)$

(ii) If f_1 and f_2 define valid inequalities, then so does $h(x) = f_1(x) f_2(x)$

(iii) For any f, the function $h(x) = f(x)^2$ defines a valid inequality

We view these as *inference rules*

Now we can use *algebra* to generate valid inequalities.

The Cone of Valid Inequalities

- The set of *polynomial* functions on $\mathbb{R}^n$ with real coefficients is denoted $\mathbb{R}[x_1, \ldots, x_n]$
- Computationally, they are easy to *parametrize*. We will consider polynomial constraint functions.

A set of polynomials $P \subset \mathbb{R}[x_1, \ldots, x_n]$ is called a *cone* if

(i) $f_1 \in P$ and $f_2 \in P$ implies $f_1 f_2 \in P$

(ii) $f_1 \in P$ and $f_2 \in P$ implies $f_1 + f_2 \in P$

(iii) $f \in \mathbb{R}[x_1, \ldots, x_n]$ implies $f^2 \in P$

It is called a *proper cone* if $-1 \notin P$

By applying the above rules to the inequality constraint functions, we can generate a *cone of valid inequalities*

Algebraic Geometry

- There is a correspondence between the *geometric object* (the feasible subset of $\mathbb{R}^n$) and the *algebraic object* (the cone of valid inequalities)
- This is a *dual* relationship; we'll see more of this later.
- The dual problem is constructed from the *cone*.
- For equality constraints, there is another algebraic object; the *ideal* generated by the equality constraints.
- For optimization, we need to look both at the geometric objects (for the primal) and the algebraic objects (for the dual problem)

Cones

- For $S \subset \mathbb{R}^n$, the cone defined by S is

$$\mathcal{C}(S) = \left\{ f \in \mathbb{R}[x_1, \dots, x_n] \;\middle|\; f(x) \geq 0 \text{ for all } x \in S \right\}$$

- If P_1 and P_2 are cones, then so is $P_1 \cap P_2$

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

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

  for some polynomials $s_1, \dots, s_r$ and some $r \geq 0$. The set of SOS polynomials Σ is a cone.

- Every cone contains Σ.

Cones

The set $\mathbf{monoid}\{f_1, \ldots, f_m\} \subset \mathbb{R}[x_1, \ldots, x_n]$ is the set of all finite products of polynomials f_i, together with 1.

The smallest cone containing the polynomials $f_1, \ldots, f_m$ is

$$\mathbf{cone}\{f_1, \ldots, f_m\} = \left\{ \sum_{i=1}^{r} s_i g_i \mid s_0, \ldots, s_r \in \Sigma, \right.$$
$$\left. g_i \in \mathbf{monoid}\{f_1, \ldots, f_m\} \right\}$$

$\mathbf{cone}\{f_1, \ldots, f_m\}$ is called the *cone generated by* $f_1, \ldots, f_m$

Explicit Parametrization of the Cone

- If $f_1, \ldots, f_m$ are valid inequalities, then so is every polynomial in $\mathbf{cone}\{f_1, \ldots, f_m\}$

- The polynomial h is an element of $\mathbf{cone}\{f_1, \ldots, f_m\}$ if and only if

$$h(x) = s_0(x) + \sum_{i=1}^{m} s_i(x) f_i(x) + \sum_{i \neq j} r_{ij}(x) f_i(x) f_j(x) + \ldots$$

where the s_i and r_{ij} are *sums-of-squares*.

An Algebraic Approach to Duality

Suppose $f_1, \ldots, f_m$ are polynomials, and consider the feasibility problem

does there exist $x \in \mathbb{R}^n$ such that

$$f_i(x) \geq 0 \qquad \text{for all } i = 1, \ldots, m$$

Every polynomial in $\mathbf{cone}\{f_1, \ldots, f_m\}$ is non-negative on the feasible set.

So if there is a polynomial $q \in \mathbf{cone}\{f_1, \ldots, f_m\}$ which satisfies

$$q(x) \leq -\varepsilon < 0 \qquad \text{for all } x \in \mathbb{R}^n$$

then the primal problem is infeasible.

Example

Let's look at the feasibility version of the previous problem. Given $t \in \mathbb{R}$, does there exist $x \in \mathbb{R}^2$ such that

$$\begin{aligned} x_1 x_2 &\leq t \\ x_1^2 + x_2^2 &\leq 1 \\ x_1 &\geq 0 \\ x_2 &\geq 0 \end{aligned}$$

Equivalently, is the set S nonempty, where

$$S = \left\{ x \in \mathbb{R}^n \mid f_i(x) \geq 0 \text{ for all } i = 1, \ldots, m \right\}$$

where

$$\begin{aligned} f_1(x) &= t - x_1 x_2 & \qquad f_2(x) &= 1 - x_1^2 - x_2^2 \\ f_3(x) &= x_1 & f_4(x) &= x_2 \end{aligned}$$

Example Continued

Let $q(x) = f_1(x) + \frac{1}{2}f_2(x)$. Then clearly $q \in \mathbf{cone}\{f_1, f_2, f_3, f_4\}$ and

$$\begin{aligned} q(x) &= t - x_1x_2 + \tfrac{1}{2}(1 - x_1^2 - x_2^2) \\ &= t + \tfrac{1}{2} - \tfrac{1}{2}(x_1 + x_2)^2 \\ &\leq t + \tfrac{1}{2} \end{aligned}$$

So for any $t < -\frac{1}{2}$, the primal problem is infeasible.

Corresponds to Lagrange multipliers $(1, 0, 0, \frac{1}{2})$ for the thm. of alternatives.

Alternatively, this is a proof by contradiction.

- If there exists x such that $f_i(x) \geq 0$ for $i = 1, \ldots, 4$ then we must also have $q(x) \geq 0$, since $q \in \mathbf{cone}\{f_1, \ldots, f_4\}$
- But we proved that q is negative if $t < -\frac{1}{2}$

Example Continued

We can also do better by using other functions in the cone. Try

$$\begin{aligned} q(x) &= f_1(x) + f_3(x) f_4(x) \\ &= t \end{aligned}$$

giving the stronger result that for any $t < 0$ the inequalities are infeasible. Again, this corresponds to Lagrange multipliers $(1, 0, 0, 1)$

- In both of these examples, we found q in the cone which was globally negative. We can view q as the Lagrangian function evaluated at a particular value of λ
- The Lagrange multiplier procedure is *searching* over a *particular subset* of functions in the cone; those which are generated by *linear combinations* of the original constraints.
- By searching over more functions in the cone we can do better

Normalization

In the above example, we have

$$q(x) = t + \tfrac{1}{2} - \tfrac{1}{2}(x_1 + x_2)^2$$

We can also show that $-1 \in \mathbf{cone}\{f_1, \ldots, f_4\}$, which gives a very simple proof of primal infeasibility.

Because, for $t < -\frac{1}{2}$, we have

$$-1 = a_0 q(x) + a_1 (x_1 + x_2)^2$$

and by construction q is in the cone, and $(x_1 + x_2)^2$ is a sum of squares.

Here a_0 and a_1 are positive constants

$$a_0 = \frac{-2}{2t+1} \qquad a_1 = \frac{-1}{2t+1}$$

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

$$f_i(x) \geq 0 \qquad \text{for all } i = 1, \ldots, m$$

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* implies that *strong duality* holds here.

Interpretation: Searching the Cone

- Lagrange duality is searching over *linear combinations* with nonnegative coefficients
$$\lambda_1 f_1 + \cdots + \lambda_m f_m$$
to find a globally negative function as a certificate

- The above algebraic procedure is searching over *conic combinations*
$$s_0(x) + \sum_{i=1}^{m} s_i(x) f_i(x) + \sum_{i \neq j} r_{ij}(x) f_i(x) f_j(x) + \ldots$$
where the s_i and r_{ij} are *sums-of-squares*

Interpretation: Formal Proof

We can also view this as a type of *formal proof*:

- View $f_1, \ldots, f_m$ are *predicates*, with $f_i(x) \geq 0$ meaning that x satisfies f_i.
- Then $\mathbf{cone}\{f_1, \ldots, f_m\}$ consists of predicates which are *logical consequences* of $f_1, \ldots, f_m$.
- If we find -1 in the cone, then we have a proof by contradiction.

Our objective is to *automatically search* the cone for negative functions; i.e., proofs of infeasibility.

Example: Linear Inequalities

Does there exist $x \in \mathbb{R}^n$ such that

$$Ax \geq 0$$

$$c^T x \leq -1$$

Write A in terms of its rows $A = \begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix}$,

then we have inequality constraints defined by linear polynomials

$$f_i(x) = a_i^T x \qquad \text{for } i = 1, \ldots, m$$

$$f_{m+1}(x) = -1 - c^T x$$

Example: Linear Inequalities

We'll search over functions $q \in \mathbf{cone}\{f_1, \ldots, f_{m+1}\}$ of the form

$$q(x) = \sum_{i=1}^{m} \lambda_i f_i(x) + \mu f_{m+1}(x)$$

Then the algebraic form of the dual is:

does there exist $\lambda_i \geq 0$, $\mu \geq 0$ such that

$$q(x) = -1 \qquad \text{for all } x$$

if the dual is feasible, then the primal problem is infeasible

Example: Linear Inequalities

The above dual condition is

$$\lambda^T A x + \mu(-1 - c^T x) = -1 \qquad \text{for all } x$$

which holds if and only if $A^T \lambda = \mu c$ and $\mu = 1$.

So we can state the duality result as follows.

Farkas Lemma

If there exists $\lambda \in \mathbb{R}^m$ such that

$$A^T \lambda = c \qquad \text{and} \qquad \lambda \geq 0$$

then there does not exist $x \in \mathbb{R}^n$ such that

$$Ax \geq 0 \qquad \text{and} \qquad c^T x \leq -1$$

Farkas Lemma

Farkas Lemma states that the following are strong alternatives

(i) there exists $\lambda \in \mathbb{R}^m$ such that $A^T\lambda = c$ and $\lambda \geq 0$

(ii) there exists $x \in \mathbb{R}^n$ such that $Ax \geq 0$ and $c^T x < 0$

Since this is just Lagrangian duality, there is a geometric interpretation

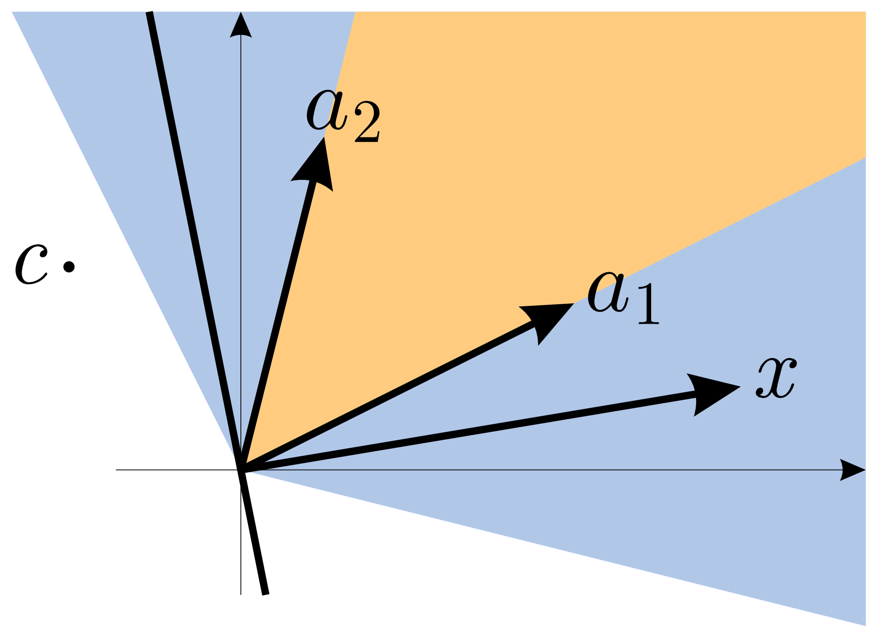
(i) c is in the convex cone

$$\{ A^T\lambda \mid \lambda \geq 0 \}$$

(ii) x defines the hyperplane

$$\{ y \in \mathbb{R}^n \mid y^T x = 0 \}$$

which separates c from the cone



Optimization Problems

Let's return to optimization problems instead of feasibility problems.

$$\begin{array}{rll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \geq 0 & \text{for all } i = 1, \ldots, m \end{array}$$

The corresponding feasibility problem is

$$\begin{aligned} t - f_0(x) &\geq 0 \\ f_i(x) &\geq 0 \qquad \text{for all } i = 1, \ldots, m \end{aligned}$$

One simple dual is to find polynomials s_i and r_{ij} such that

$$t - f_0(x) + \sum_{i=1}^{m} s_i(x) f_i(x) + \sum_{i \neq j} r_{ij}(x) f_i(x) f_j(x) + \ldots$$

is globally negative, where the s_i and r_{ij} are *sums-of-squares*

Optimization Problems

We can combine this with a maximization over t

$$\begin{aligned} \text{maximize} \quad & t \\ \text{subject to} \quad & t - f_0(x) + \sum_{i=1}^{m} s_i(x) f_i(x) + \\ & \qquad \sum_{i=1}^{m} \sum_{j=1}^{m} r_{ij}(x) f_i(x) f_j(x) \le 0 \text{ for all } x \\ & s_i, r_{ij} \text{ are sums-of-squares} \end{aligned}$$

- The variables here are (coefficients of) the polynomials s_i, r_i
- We will see later how to approach this kind of problem using semidefinite programming