Global minimization of rational functions using semidefinite programming

Etienne de Klerk † , Dorina Jibetean ‡

[†]Delft University of Technology, [‡]CWI, Amsterdam

Rational function minimization

Let $p, q, p_1, \ldots, p_k \in \mathbb{R}[x_1, \ldots, x_n]$ (polynomials with real coefficients defined on \mathbb{R}^n) with p and q relatively prime.

Rational function minimization

Let $p, q, p_1, \ldots, p_k \in \mathbb{R}[x_1, \ldots, x_n]$ (polynomials with real coefficients defined on \mathbb{R}^n) with p and q relatively prime. We will consider the optimization problem:

$$p^* := \inf_{x \in S} \frac{p(x)}{q(x)}$$

where *S* is the *semi-algebraic set* given by

 $S := \{ x \in \mathbb{R}^n : p_i(x) \ge 0, \ i = 1, \dots, k \}.$

Rational function minimization

Let $p, q, p_1, \ldots, p_k \in \mathbb{R}[x_1, \ldots, x_n]$ (polynomials with real coefficients defined on \mathbb{R}^n) with p and q relatively prime. We will consider the optimization problem:

$$p^* := \inf_{x \in S} \frac{p(x)}{q(x)}$$

where *S* is the *semi-algebraic set* given by

 $S := \{ x \in \mathbb{R}^n : p_i(x) \ge 0, \ i = 1, \dots, k \}.$

 p^* is not necessarily attained or finite!

Applications

• Least squares approximation of data using rational functions (least squares Padé approximation);

Applications

- Least squares approximation of data using rational functions (least squares Padé approximation);
- *H*₂ model reduction (D. Jibetean. PhD Thesis, CWI, Amsterdam, 2003.);

Applications

- Least squares approximation of data using rational functions (least squares Padé approximation);
- *H*₂ model reduction (D. Jibetean. PhD Thesis, CWI, Amsterdam, 2003.);
- stability analysis of certain dynamical systems, including biochemical reactor models.

Possible approaches

• If the infimum is attained one can solve the first order optimality condition equations. Modern review: B. Sturmfels, *Solving Systems of Polynomial Equations*, AMS, 2002. If the inf is not attained ...

Possible approaches

- If the infimum is attained one can solve the first order optimality condition equations. Modern review: B. Sturmfels, *Solving Systems of Polynomial Equations*, AMS, 2002. If the inf is not attained ...
- Global optimization codes can converge to local minima.

Possible approaches

- If the infimum is attained one can solve the first order optimality condition equations. Modern review: B. Sturmfels, *Solving Systems of Polynomial Equations*, AMS, 2002. If the inf is not attained ...
- Global optimization codes can converge to local minima.
- Today's talk: approaches involving semidefinite programming (SDP).

SDP is a generalization of LP and interior point algorithms for LP can be extended to SDP.

SDP is a generalization of LP and interior point algorithms for LP can be extended to SDP.



SDP is a generalization of LP and interior point algorithms for LP can be extended to SDP.



subject to
$$\sum_{i,j=1}^{n} a_{ij}^{(k)} x_{ij} = b_k \quad \forall \ k = 1, ..., m,$$

SDP is a generalization of LP and interior point algorithms for LP can be extended to SDP.

$$\inf_{[x_{ij}]} \sum_{i,j=1}^{n} c_{ij} x_{ij}$$

subject to
$$\sum_{i,j=1}^{n} a_{ij}^{(k)} x_{ij} = b_k \quad \forall \ k = 1, ..., m,$$

 $X := [x_{ij}] \succeq 0 \text{ (p.s.d.)}$

SDP is a generalization of LP and interior point algorithms for LP can be extended to SDP.

$$\inf_{[x_{ij}]} \sum_{i,j=1}^{n} c_{ij} x_{ij}$$

subject to
$$\sum_{i,j=1}^{n} a_{ij}^{(k)} x_{ij} = b_k \quad \forall \ k = 1, ..., m,$$

 $X := [x_{ij}] \succeq 0 \text{ (p.s.d.)}$

If the data matrices diagonal \Rightarrow LP

Different cases

We investigate SDP-based approaches for the following cases of $\inf_{x \in S} p(x)/q(x)$:

• $S = \mathbb{R}^n$ and n = 1 (Unconstrained minimization: univariate case);

Different cases

We investigate SDP-based approaches for the following cases of $\inf_{x \in S} p(x)/q(x)$:

- $S = \mathbb{R}^n$ and n = 1 (Unconstrained minimization: univariate case);
- $S = \mathbb{R}^n$ and general *n* (Unconstrained minimization: general case);

Different cases

We investigate SDP-based approaches for the following cases of $\inf_{x \in S} p(x)/q(x)$:

- $S = \mathbb{R}^n$ and n = 1 (Unconstrained minimization: univariate case);
- $S = \mathbb{R}^n$ and general *n* (Unconstrained minimization: general case);
- *S* is compact, connected and general *n* (Constrained case);

Unconstrained case

Consider the unconstrained problem.

$$p^* := \inf_{x \in \mathbb{R}^n} \frac{p(x)}{q(x)}$$
$$= \sup \left\{ \rho : \frac{p(x)}{q(x)} - \rho \ge 0 \quad \forall x \in \mathbb{R}^n \right\}$$

Unconstrained case

Consider the unconstrained problem.

$$p^* := \inf_{x \in \mathbb{R}^n} \frac{p(x)}{q(x)}$$
$$= \sup \left\{ \rho : \frac{p(x)}{q(x)} - \rho \ge 0 \quad \forall x \in \mathbb{R}^n \right\}$$

We can replace the nonnegativity condition by a simpler one ...

Theorem (Jibetean) Assume $p^* > -\infty$. Then *q does not change sign* on \mathbb{R}^n .

Theorem (Jibetean) Assume $p^* > -\infty$. Then *q* does not change sign on \mathbb{R}^n . Assuming $q(x) \ge 0$ on *S* one has:

$$\frac{p(x)}{q(x)} - \rho \ge 0 \quad \forall x \in \mathbb{R}^n$$

iff

 $p(x) - \rho q(x) \ge 0 \quad \forall x \in \mathbb{R}^n.$

Theorem (Jibetean) Assume $p^* > -\infty$. Then *q* does not change sign on \mathbb{R}^n . Assuming $q(x) \ge 0$ on *S* one has:

$$\frac{p(x)}{q(x)} - \rho \ge 0 \quad \forall x \in \mathbb{R}^n$$

iff

$$p(x) - \rho q(x) \ge 0 \quad \forall x \in \mathbb{R}^n.$$

D. Jibetean. Global optimization of rational multivariate functions. Technical Report PNA-R0120, CWI, Amsterdam, 2001.

Theorem (Jibetean) Assume $p^* > -\infty$. Then *q* does not change sign on \mathbb{R}^n . Assuming $q(x) \ge 0$ on *S* one has:

$$\frac{p(x)}{q(x)} - \rho \ge 0 \quad \forall x \in \mathbb{R}^n$$

iff

$$p(x) - \rho q(x) \ge 0 \quad \forall x \in \mathbb{R}^n.$$

D. Jibetean. Global optimization of rational multivariate functions. Technical Report PNA-R0120, CWI, Amsterdam, 2001.

This leads us to the theory of *nonnegative polynomials*.

Let $p \in \mathbb{R}[x_1, \ldots, x_n]$.

Let $p \in \mathbb{R}[x_1, \ldots, x_n]$.

p is called a *sum of squares* (SOS) if there exist polynomials p_i such that $p = \sum_i p_i^2$.

Let $p \in \mathbb{R}[x_1, \ldots, x_n]$.

p is called a *sum of squares* (SOS) if there exist polynomials p_i such that $p = \sum_i p_i^2$.

Nonnegativity and *sum of squares* are the same if:

Let $p \in \mathbb{R}[x_1, \ldots, x_n]$.

p is called a *sum of squares* (SOS) if there exist polynomials p_i such that $p = \sum_i p_i^2$.

Nonnegativity and *sum of squares* are the same if:

n = 1 (univariate polynomials) (result by Markov?);

Let $p \in \mathbb{R}[x_1, \ldots, x_n]$.

p is called a *sum of squares* (SOS) if there exist polynomials p_i such that $p = \sum_i p_i^2$.

Nonnegativity and *sum of squares* are the same if:

- n = 1 (univariate polynomials) (result by Markov?);
- d = 2 (quadratic polynomials on *n* variables);

Let $p \in \mathbb{R}[x_1, \ldots, x_n]$.

p is called a *sum of squares* (SOS) if there exist polynomials p_i such that $p = \sum_i p_i^2$.

Nonnegativity and *sum of squares* are the same if:

- n = 1 (univariate polynomials) (result by Markov?);
- d = 2 (quadratic polynomials on *n* variables);
- n = 2 and $d \le 4$ (bivariate polynomials of degree at most 4) (result by Hilbert);

Let $p \in \mathbb{R}[x_1, \ldots, x_n]$.

p is called a *sum of squares* (SOS) if there exist polynomials p_i such that $p = \sum_i p_i^2$.

Nonnegativity and *sum of squares* are the same if:

- n = 1 (univariate polynomials) (result by Markov?);
- d = 2 (quadratic polynomials on *n* variables);
- n = 2 and $d \le 4$ (bivariate polynomials of degree at most 4) (result by Hilbert);

In all other cases counterexamples exist.

The sum of squares cone

We fix a basis of monomials

 $\tilde{x}_{n,d} := (1, x_1, \dots, x_n, x_1^2, \dots, x_n^d) \dim \binom{n+d}{d}.$

The sum of squares cone

We fix a basis of monomials

$$\tilde{x}_{n,d} := (1, x_1, \dots, x_n, x_1^2, \dots, x_n^d) \operatorname{dim:} \binom{n+d}{d}.$$

Notation: We denote the convex cone generated by squares of polynomials on \mathbb{R}^n of degree at most *d* by $\sum_{n,2d}^2 (sum-of-squares (SOS) cone).$

The sum of squares cone

We fix a basis of monomials

$$\tilde{x}_{n,d} := (1, x_1, \dots, x_n, x_1^2, \dots, x_n^d) \operatorname{dim:} \binom{n+d}{d}.$$

Notation: We denote the convex cone generated by squares of polynomials on \mathbb{R}^n of degree at most *d* by $\sum_{n,2d}^2 (sum-of-squares (SOS) cone).$

(We drop the subscripts when they are clear from the context.)

The sum of squares cone (cdt.)

Theorem: For a given $f \in \mathbb{R}[x_1, \dots, x_n]$ of degree 2*d*, one has $f \in \sum_{n,2d}^2$ iff

$$f = \tilde{x}_{n,d}^T M \tilde{x}_{n,d}$$

for some
$$M \succeq 0$$
 (size $\binom{n+d}{d} \times \binom{n+d}{d}$).

The sum of squares cone (cdt.)

Theorem: For a given $f \in \mathbb{R}[x_1, \dots, x_n]$ of degree 2*d*, one has $f \in \sum_{n,2d}^2$ iff

$$f = \tilde{x}_{n,d}^T M \tilde{x}_{n,d}$$

for some $M \succeq 0$ (size $\binom{n+d}{d} \times \binom{n+d}{d}$).

Implication: Conic linear optimization over the cone $\sum_{n,d}^2$ can be done using *semidefinite programming* (SDP);
The sum of squares cone (cdt.)

Theorem: For a given $f \in \mathbb{R}[x_1, \dots, x_n]$ of degree 2*d*, one has $f \in \sum_{n,2d}^2$ iff

$$f = \tilde{x}_{n,d}^T M \tilde{x}_{n,d}$$

for some $M \succeq 0$ (size $\binom{n+d}{d} \times \binom{n+d}{d}$).

Implication: Conic linear optimization over the cone $\sum_{n,d}^2$ can be done using *semidefinite programming* (SDP);

cf. Theorem 17.1 in Y. Nesterov. Squared functional systems and optimization problems. In J.B.G. Frenk et al. eds., *High performance optimization*, 405–440. KAP, 2000.

If q is nonnegative on \mathbb{R} , then

$$\inf_{x \in \mathbb{R}} \frac{p(x)}{q(x)} = \sup_{t,x} \{t : p(x) - tq(x) \ge 0 \ \forall x \in \mathbb{R} \}$$
$$= \sup_{t,x} \{t : p(x) - tq(x) \in \Sigma^2 \}$$
$$= \sup_{t,x} \{t : p(x) - tq(x) = \tilde{x}^T M \tilde{x} \}$$

for some $M \succeq 0$, where

 $\tilde{x}^T = [1 \ x \ x^2 \dots x^{\frac{1}{2} \max\{\deg(p), \deg(q)\}}].$

Let $p(x) - tq(x) = \sum_{\alpha} a_{\alpha}(t) x^{\alpha}$. NB: $a_{\alpha}(t)$ is affine in t.

Let $p(x) - tq(x) = \sum_{\alpha} a_{\alpha}(t)x^{\alpha}$. NB: $a_{\alpha}(t)$ is affine in t. Then the optimization problem becomes: maximize t such that

$$a_{\alpha}(t) = \sum_{i+j=\alpha} M_{ij}, \quad M \succeq 0.$$

This is an SDP problem!

Let $p(x) - tq(x) = \sum_{\alpha} a_{\alpha}(t)x^{\alpha}$. NB: $a_{\alpha}(t)$ is affine in t. Then the optimization problem becomes: maximize t such that

$$a_{\alpha}(t) = \sum_{i+j=\alpha} M_{ij}, \quad M \succeq 0.$$

This is an SDP problem! (Result already obtained by Nesterov for $q(x) \equiv 1$.)

Y. Nesterov. Squared functional systems and optimization problems. In J.B.G. Frenk et al. eds., *High performance optimization*, 405–440. KAP, 2000.

Example

 $\frac{p(x)}{q(x)} := \frac{x^2 - 2x}{(x+1)^2}.$



Global minimization of rational functions using semidefinite programming $-\,p.14/24$

$$\frac{p(x)}{q(x)} := \frac{x^2 - 2x}{(x+1)^2}.$$

$$\frac{p(x)}{q(x)} := \frac{x^2 - 2x}{(x+1)^2}$$

Equivalent problem: $\sup t$ such that

$$(1-t)x^{2}-2(1+t)x-t = \begin{bmatrix} 1 \\ x \end{bmatrix}^{T} \begin{bmatrix} M_{00} & M_{01} \\ M_{10} & M_{11} \end{bmatrix} \begin{bmatrix} 1 \\ x \end{bmatrix},$$
(2)

for some $M \succeq 0$.

From (2):

 $M_{00} = -t$, $M_{01} = M_{10} = -(1+t)$, $M_{11} = 1-t$.

From (2):

 $M_{00} = -t$, $M_{01} = M_{10} = -(1+t)$, $M_{11} = 1-t$.

We therefore get

$$\min_{x \in \mathbb{R}} \frac{p(x)}{q(x)} = \max_{t, M} t$$

such that

$$M = \begin{bmatrix} -t & -(1+t) \\ -(1+t) & 1-t \end{bmatrix} \succeq 0.$$

From (2):

 $M_{00} = -t$, $M_{01} = M_{10} = -(1+t)$, $M_{11} = 1-t$.

We therefore get

$$\min_{x \in \mathbb{R}} \frac{p(x)}{q(x)} = \max_{t, M} t$$

such that

$$M = \begin{bmatrix} -t & -(1+t) \\ -(1+t) & 1-t \end{bmatrix} \succeq 0.$$

Note that the optimal value is $p^* = -1/3$.

Consider a *semi-algebraic set*

 $S = \{x \in \mathbb{R}^n : p_i(x) \ge 0 \ (i = 1, \dots, k)\}.$

Consider a *semi-algebraic set*

 $S = \{x \in \mathbb{R}^n : p_i(x) \ge 0 \ (i = 1, \dots, k)\}.$

General constrained problem: find

$$p^* \coloneqq \inf_{x \in S} \frac{p(x)}{q(x)}.$$

Consider a *semi-algebraic set*

 $S = \{x \in \mathbb{R}^n : p_i(x) \ge 0 \ (i = 1, \dots, k)\}.$

General constrained problem: find

$$p^* \coloneqq \inf_{x \in S} \frac{p(x)}{q(x)}.$$

One can treat the unconstrained multivariate problem by adding an artificial constraint $||x||^2 \leq R$ for some 'large' R.

Theorem (Jibetean) Assume that *S* is open and connected (or the (partial) closure of such a set). If $p^* > -\infty$ then *q* does not change sign on *S*. Assuming $q(x) \ge 0$ on *S*, then

 $\frac{p(x)}{q(x)} \ge \alpha \; \forall x \in S \iff p(x) - \alpha q(x) \ge 0 \; \forall x \in S.$

Theorem (Jibetean) Assume that *S* is open and connected (or the (partial) closure of such a set). If $p^* > -\infty$ then *q* does not change sign on *S*. Assuming $q(x) \ge 0$ on *S*, then

 $\frac{p(x)}{q(x)} \ge \alpha \; \forall x \in S \iff p(x) - \alpha q(x) \ge 0 \; \forall x \in S.$

D. Jibetean. PhD Thesis, CWI, Amsterdam, 2003.

Theorem (Jibetean) Assume that *S* is open and connected (or the (partial) closure of such a set). If $p^* > -\infty$ then *q* does not change sign on *S*. Assuming $q(x) \ge 0$ on *S*, then

$$\frac{p(x)}{q(x)} \ge \alpha \; \forall x \in S \iff p(x) - \alpha q(x) \ge 0 \; \forall x \in S.$$

D. Jibetean. PhD Thesis, CWI, Amsterdam, 2003.

Consequence

$$\inf_{x \in S} \frac{p(x)}{q(x)} = \sup \left\{ \rho : p(x) - \rho q(x) \ge 0 \ \forall x \in S \right\}.$$

Global minimization of rational functions using semidefinite programming -p.18/24

Constrained multivariate case Technical assumption: S is compact and there exists a

 $\bar{p} \in \Sigma^2 + p_1 \Sigma^2 + \ldots + p_k \Sigma^2$

such that $\{x : \overline{p}(x) \ge 0\}$ is *compact*.

Constrained multivariate case Technical assumption: *S* is compact and there exists a $\bar{p} \in \Sigma^2 + p_1 \Sigma^2 + \ldots + p_k \Sigma^2$

such that $\{x : \overline{p}(x) \ge 0\}$ is *compact*.

Theorem (Putinar): For a given polynomial p_0 one has $p_0(x) > 0$ for all $x \in S$ iff

$$p_0 \in \Sigma^2 + p_1 \Sigma^2 + \ldots + p_k \Sigma^2.$$

Constrained multivariate case Technical assumption: *S* is compact and there exists a $\bar{p} \in \Sigma^2 + p_1 \Sigma^2 + \ldots + p_k \Sigma^2$

such that $\{x : \overline{p}(x) \ge 0\}$ is *compact*.

Theorem (Putinar): For a given polynomial p_0 one has $p_0(x) > 0$ for all $x \in S$ iff

$$p_0 \in \Sigma^2 + p_1 \Sigma^2 + \ldots + p_k \Sigma^2.$$

M. Putinar. Positive polynomials on compact semi-algebraic sets. *Ind. Univ. Math. J.* 42:969–984, 1993.

Consider the minimization problem

$$p^* = \inf_{x \in S} \frac{p(x)}{q(x)}.$$

Consider the minimization problem

$$p^* = \inf_{x \in S} \frac{p(x)}{q(x)}.$$

If p and q have no common roots in S, then by Putinar's and Jibetean's theorems:

 $p^* = \sup \{ \rho : p(x) - \rho q(x) > 0 \ \forall x \in S \}$

 $= \sup \left\{ \rho : (p - \rho q) \in \Sigma^2 + p_1 \Sigma^2 + \ldots + p_k \Sigma^2 \right\}$

 $\geq \sup \left\{ \rho : (p - \rho q) \in \Sigma_{n,t}^2 + p_1 \Sigma_{n,t}^2 + \ldots + p_k \Sigma_{n,t}^2 \right\}$

:= ρ_t (for any integer $t \ge 1$).

We have that $\rho_i \leq \rho_{i+1} \leq p^*$ and - if p and q have no common roots in S –

$$\lim_{t \to \infty} \rho_t = p^*.$$

We have that $\rho_i \leq \rho_{i+1} \leq p^*$ and - if p and q have no common roots in S –

$$\lim_{t \to \infty} \rho_t = p^*.$$

Computation of ρ_t : SDP problem with matrices of size $\binom{n+t}{t} \times \binom{n+t}{t}$ and at most $\max\{\deg(p), \deg(q)\}$ constraints — "polynomial" complexity for t = O(1).

We have that $\rho_i \leq \rho_{i+1} \leq p^*$ and - if p and q have no common roots in S –

$$\lim_{t \to \infty} \rho_t = p^*.$$

Computation of ρ_t : SDP problem with matrices of size $\binom{n+t}{t} \times \binom{n+t}{t}$ and at most $\max\{\deg(p), \deg(q)\}$ constraints — "polynomial" complexity for t = O(1).

These results by already obtained by Lasserre for $q(x) \equiv 1$ (polynomial objective function).

J.B. Lasserre. Global optimization with polynomials and the problem of moments. *SIOPT*, 11:296–817, 2001. Global minimization of rational functions using semidefinite programming – p.21/24

Return to the unconstrained case



Return to the unconstrained case



Artificial constraint $||x||^2 \leq R$ for some 'sufficiently large' R.

Return to the unconstrained case

$$\inf_{x \in \mathbb{R}^n} \frac{p(x)}{q(x)}.$$

Artificial constraint $||x||^2 \leq R$ for some 'sufficiently large' R. Now we have $\min_{x \in S} \frac{p(x)}{q(x)}$ where S is the compact semi-algebraic set

$$S := \left\{ x \in \mathbb{R}^n : R - ||x||^2 \ge 0 \right\}.$$

Return to the unconstrained case

$$\inf_{x \in \mathbb{R}^n} \frac{p(x)}{q(x)}.$$

Artificial constraint $||x||^2 \le R$ for some 'sufficiently large' R. Now we have $\min_{x \in S} \frac{p(x)}{q(x)}$ where S is the compact semi-algebraic set

$$S := \left\{ x \in \mathbb{R}^n : R - ||x||^2 \ge 0 \right\}.$$

No a priori choice for R available in general.

• Lasserre's approach implemented in the software *GloptiPoly*.

- Lasserre's approach implemented in the software *GloptiPoly*.
- Optimization over Σ^2 implemented in *SOStools* by Parrilo et al.

- Lasserre's approach implemented in the software *GloptiPoly*.
- Optimization over Σ^2 implemented in *SOStools* by Parrilo et al.

These are add-on routines for the SDP solver *SeDuMi* by Sturm.

- Lasserre's approach implemented in the software *GloptiPoly*.
- Optimization over Σ^2 implemented in *SOStools* by Parrilo et al.

These are add-on routines for the SDP solver *SeDuMi* by Sturm. All freely available via Helmberg's SDP page:

http://www-user.tu-chemnitz.de/~helmberg/semidef.html

- Lasserre's approach implemented in the software *GloptiPoly*.
- Optimization over Σ^2 implemented in *SOStools* by Parrilo et al.

These are add-on routines for the SDP solver *SeDuMi* by Sturm. All freely available via Helmberg's SDP page:

http://www-user.tu-chemnitz.de/~helmberg/semidef.html

GloptiPoly and SOStools extremely useful to prove *global optimality* in small problems.

Discussion

• We have extended results by Nesterov, Lasserre and Parrilo to include rational objective functions.

Discussion

- We have extended results by Nesterov, Lasserre and Parrilo to include rational objective functions.
- Techniques from real algebraic geometry available to compute all KKT points, but SDP approach computationally attractive.
 See: P. Parrilo and B. Sturmfels. Minimizing polynomial

functions, 2001. (Available at arXiv.org e-Print archive)
Discussion

- We have extended results by Nesterov, Lasserre and Parrilo to include rational objective functions.
- Techniques from real algebraic geometry available to compute all KKT points, but SDP approach computationally attractive.
 See: P. Parrilo and B. Sturmfels. Minimizing polynomial

functions, 2001. (Available at arXiv.org e-Print archive)

• SDP approach competitive with state-of-the-art global optimization software.

Discussion

- We have extended results by Nesterov, Lasserre and Parrilo to include rational objective functions.
- Techniques from real algebraic geometry available to compute all KKT points, but SDP approach computationally attractive.
 See: P. Parrilo and B. Sturmfels. Minimizing polynomial
 - functions, 2001. (Available at arXiv.org e-Print archive)
- SDP approach competitive with state-of-the-art global optimization software.
- Need for large-scale (parallel?) SDP solvers to solve the large SDP relaxations.