Global Optimization of Rational Multivariate Functions

Dorina Jibetean
Centrum voor Wiskunde en Informatica
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
d.jibetean@cwi.nl

ABSTRACT
The paper deals with unconstrained global minimization of rational functions. A necessary condition is given for the function to have a finite infimum. In case the condition is satisfied, the problem is shown to be equivalent to a specific constrained polynomial optimization problem. In this paper, we solve a relaxation of the latter formulation using semi-definite programming. In general, the relaxation will produce a lower bound of the infimum. However, under no degeneracies, it is possible to check whether the relaxation was in fact exact.

2000 Mathematics Subject Classification: 90C30, 14P99, 90C22.
Keywords and Phrases: global optimization, rational functions, real algebraic geometry, sums of squares, semi-definite programming.
Note: Research is supported by the NWO Project *Constructive Algebra for System Identification and Model Reduction.*

1 Introduction

Finding the infimum on $\mathbf{R}^n$ of an arbitrary function is, to the best of our knowledge, an open problem. Numerical algorithms used to solve such problems do not give any guarantees that a global optimum is obtained. Other global approaches are based on finding all critical values in order to determine the smallest one, but of course they work only in case the function has a minimum (i.e., the infimum is attained).

So far, the class of polynomial functions is the only class for which algorithms guaranteed to find the infimum have been developed. Here, I'll mention only those approaches which do not assume the existence of a minimum, hence which do not use the first order conditions. There are two such different approaches to the polynomial optimization problem.
One is based on the relation between positive polynomials and sums of squares of polynomials. In fact, in this approach the effort is directed towards finding a real number α such that $p(x) - \alpha \geq 0$, $\forall x \in \mathbf{R}^n$, where p is the given polynomial. The algorithm finds in general a lower bound of the polynomial's infimum. However, under no degeneracies, it can be checked if this is the true infimum or not. The method is presented in [3].
A completely different method for computing the infimum of a polynomial function is given in [2]. There, a particular perturbation of the polynomial allows one to find the infimum in the original problem by looking at the minimum of the perturbed polynomial. The perturbation is done such that the perturbed polynomial has a minimum, and therefore the first order conditions of the perturbed polynomial can be used. The second method has a theoretical advantage over the first one mentioned since it always returns the infimum (and not a lower bound of it). However, from the

computational point of view, it seems to be more demanding.

In this paper we are going to deal with the more general case of optimization of rational functions over $\mathbf{R}^n$. Excepting some approaches based on solving the associated first order conditions (and which therefore assume that the minimum exists), the problem did not receive particular attention so far.

Our approach extends the first method mentioned, designed for polynomial optimization. Hence, our algorithm will return in general a lower bound of the actual infimum, with the possibility of checking in some cases, as discussed, whether that equals the actual infimum or not. An interesting criterion will be given for a rational function to have the infimum at $-\infty$.

The paper is organized as follows: Section 2 presents a preliminary, general result which is also the main theoretical result of this paper. From this, we deduce in Section 3 necessary conditions for a rational function to be bounded from below. Moreover, we give a theorem which translates in an equivalent way the rational optimization problem into a specific constrained polynomial optimization problem. For the new constrained polynomial optimization problem we show the applicability of the method developed by [4], [3] which reduces it to a semi-definite programming problem. This is the subject of Section 4. An example is presented in Section 5 and conclusions are drawn in the last section.

2 Preliminary result

Theorem 2.1 *Let $a(x)/b(x)$ be a rational multivariate function, with $a(x)$, $b(x)$ relatively prime polynomials. If $a(x)/b(x) \geq 0$, $\forall x \in \mathbf{R}^n \setminus \{x \in \mathbf{R}^n \mid b(x) = 0\}$, then one of the two following statements holds:*

- $a(x) \geq 0, \quad b(x) \geq 0 \quad \forall x \in \mathbf{R}^n$,
- $a(x) \leq 0, \quad b(x) \leq 0 \quad \forall x \in \mathbf{R}^n$.

Proof
Note that the condition $a(x)/b(x) \geq 0$, $\forall x \in \mathbf{R}^n \setminus \{x \in \mathbf{R}^n \mid b(x) = 0\}$ is equivalent, by multiplication with $b^2(x)$, to $a(x)b(x) \geq 0 \ \forall x \in \mathbf{R}^n$.
The proof of the theorem is based on showing that the decomposition of the polynomial $a(x)b(x)$ into irreducible factors has the following form

$$a(x)b(x) = \prod_{i=1}^{K_1} g_i(x)^{2m_i} \prod_{j=K_1+1}^{K_2} g_j(x)^{m_j},$$

where $g_i, i = 1, \ldots, K_1$ are all the factors that change sign on $\mathbf{R}^n$ and $g_j, j = K_1 + 1, \ldots, K_2$ the factors that do not change sign on $\mathbf{R}^n$. In other words, we prove that if there exists an irreducible divisor of $a(x)b(x)$ that changes sign on $\mathbf{R}^n$, it actually has an even power in the decomposition of $a(x)b(x)$.
Using the decomposition above and the fact that a, b are relatively prime polynomials, it's clear that neither a nor b changes sign on $\mathbf{R}^n$. Since their product is non-negative, it also implies that in fact they are both either non-negative or non-positive.
Let us consider $g_1 \in \mathbf{R}[x]$, an irreducible divisor of $g(x) = a(x)b(x)$ which changes sign. By Theorem 4.5.1 of [1], the ideal generated by g_1 is a real ideal.
Let us denote

$$\frac{g(x)}{g_1(x)} = \tilde{g_1}(x),$$

which can be rewritten equivalently

$$g(x) = \tilde{g_1}(x)g_1(x) \geq 0 \quad \forall x \in \mathbf{R}^n.$$

Hence there exists the polynomials $r(x),\ s_i(x),\ i = 1, \ldots, m$ ([1], Theorem 6.1.1) such that

$$r^2(x)\tilde{g_1}(x)g_1(x) = \sum_{i=1}^{m} s_i^2(x) \quad \forall x \in \mathbf{R}^n. \tag{1}$$

Take r minimal with respect to the division, having the property that $r^2(x)\tilde{g_1}(x)g_1(x)$ can be written as a sum of squares of polynomials.
The left hand side obviously belongs to the real ideal (g_1). By the definition of a real ideal ([1], Definition 4.1.3), the above relation implies that $s_i \in (g_1), \quad \forall i = 1, \ldots, n$. Hence there exist polynomials $t_i(x)$ such that $s_i(x) = t_i(x)g_1(x)$.
By replacing s_i's in (1) and dividing both sides of the equality by g_1 we get

$$r^2(x)\tilde{g_1}(x) = g_1(x)\sum_{i=1}^{m} t_i^2(x) \quad \forall x \in \mathbf{R}^n. \tag{2}$$

Therefore g_1 must divide $r^2(x)\tilde{g_1}(x)$ and since g_1 is irreducible, g_1 divides $\tilde{g_1}$ or g_1 divides r.
Suppose first that g_1 divides r. Then there exists a polynomial $r_1(x)$ satisfying $r(x) = g_1(x)r_1(x)$.
By replacing r into (2) and dividing both sides of the equality by $g_1(x)$ we obtain

$$r_1^2(x)\tilde{g_1}(x)g_1(x) = \sum_{i=n}^{m} t_i^2(x), \quad \forall x \in \mathbf{R}^n.$$

However, by comparing with (1) we obtain a contradiction with the minimality of r.
Hence it must be that g_1 divides $\tilde{g_1}$ which implies g_1^2 divides g.
By applying exactly the same procedure to the polynomial g/g_1^2, one can show that any irreducible factor of $g(x) = a(x)b(x)$ which changes sign must have an even power in the decomposition of $g(x)$.
This concludes the proof. □

3 Application to rational optimization problems

Consider the following problem

$$\inf_{x \in \mathbf{R}^n} \frac{p(x)}{q(x)}, \quad \text{with} \quad p(x),\ q(x) \in \mathbf{R}[x] \quad \text{relatively prime.} \tag{3}$$

Regarding the terminology, we are using *infimum* (inf) instead of the more common *minimum* (min) or, later on, *supremum* (sup) instead of *maximum* (max) simply to stress that the optimal value may not be attained in $\mathbf{R}^n$ but only approached asymptotically. Note that there are no other differences between the formulations involving inf (respectively sup) and min (respectively max).

In the following we use Theorem 2.1 of the previous section in order to obtain some criteria for our problem.

Proposition 3.1 *Let $p(x)/q(x)$ be a rational function with $p(x),\ q(x)$ relatively prime. If $p(x)/q(x)$ is bounded from below, then q has constant sign on $\mathbf{R}^n$.*

Proof
Let $\alpha \in \mathbf{R}$ such that $p(x)/q(x) \geq \alpha \ \ \forall x \in \mathbf{R}^n$. Then $(p(x) - \alpha q(x))/q(x)$ satisfies the hypothesis of Theorem 2.1 hence $p(x) - \alpha q(x)$ and $q(x)$ have constant sign on $\mathbf{R}^n$. □

An immediate consequence is formulated below:

Corollary 3.2 *Let $p(x)/q(x)$ be a rational function with $p(x)$, $q(x)$ relatively prime polynomials. If $q(x)$ changes sign on $\mathbf{R}^n$ then* $\inf_{x\in\mathbf{R}^n} p(x)/q(x) = -\infty$.

Note that the reciprocal is not true.

However, we can reformulate now the problem (3). Suppose that $q(x) \geq 0 \quad \forall x \in \mathbf{R}^n$. Then using the proof of Proposition 3.1 and Theorem 2.1, problem (3) is equivalent to

$$\begin{array}{ll} \sup & \alpha \\ \text{s.t.} & p(x) - \alpha q(x) \geq 0, \quad \forall x \in \mathbf{R}^n. \end{array} \tag{4}$$

Obviously the largest α satisfying the condition is the infimum of $p(x)/q(x)$.

Note that the feasibility domain of (4) may be the empty set. That is, there is no $\alpha \in \mathbf{R}$ satisfying the polynomial inequality for every $x \in \mathbf{R}^n$. In this case the supremum will be $-\infty$.

The condition $q(x) \geq 0 \quad \forall x \in \mathbf{R}^n$ can be checked in the following way. Evaluate q at an arbitrary point and suppose that it is indeed positive. Then q is non-negative on $\mathbf{R}^n$ if and only if $\inf_{x\in\mathbf{R}^n} q(x) \geq 0$. Hence we only need to compute the infimum of a polynomial on $\mathbf{R}^n$ and this can be done using for example the algorithm described in [2].

To conclude, in this section we have rewritten the rational optimization problem as a constrained polynomial optimization problem. Several options are possible now. One of them is discussed in the next section.

4 A semi-definite programming relaxation

In this section we study the extension to rational functions of a method based on [4] and [3], used previously for polynomial functions. As in [3], we want to rewrite the rational optimization problem into a semi-definite optimization problem (SDP) which is known to have good computational complexity. Actually, in general we obtain an SDP relaxation of the original problem, which gives a lower bound for the solution of the original problem.

Let us study now how to rewrite the problem (4) as an SDP. For this we study the polynomial $p(x) - \alpha q(x)$.

Let us denote $F(x) = p(x) - \alpha q(x)$. If the total degree of F is odd, then its infimum will be $-\infty$ (take all variables equal). Hence, in this case the polynomial cannot be positive everywhere. Then (3) is $-\infty$. It is therefore sufficient to restrict ourselves to the case of even degree polynomials. We follow closely [3] and produce a relaxation of the problem (4). Let F have a total degree $2d$. We want to find a matrix Q such that

$$F(x) = z^T Q z, \quad , z = [1, x_1, x_2, \ldots, x_n, x_1 x_2, \ldots, x_n^d].$$

z contains all monomials in the variables $x_1, \ldots, x_n$ of degree less than or equal to d. Obviously, if such Q exists, then it is a symmetric matrix.

The way to construct such a matrix Q is described in the proof below, where we also argue that Q can always be constructed. If $Q \succeq 0$, then $F(x) \geq 0$, $\forall x \in \mathbf{R}^n$. Conversely, it is not always true. There exist examples of polynomials which are positive on $\mathbf{R}^n$ and for which no matrix Q, constructed as described above, is positive semi-definite. This results are very closely related to Hilbert 17'th problem (see [1]). This is the reason for which our reformulation is in general just a relaxation and not equivalent to the initial rational optimization problem.

Note also that Q is not necessarily unique. In [3] it is shown that Q belongs to an affine subspace, parameterized by λ's. We extend this result to the case of $F(x) = p(x) - \alpha q(x)$ (hence depending on α) and show that the generic matrix Q is affine in α and λ's.

Theorem 4.1 *Let the symmetric matrix Q satisfy $p(x) - \alpha q(x) = z^T Q z$. Then Q belongs to an affine subspace.*

Proof
This is a constructive proof. We compute the matrix Q by making the computation on the right-hand side of $p(x) - \alpha q(x) = z^T Q z$, where z depends entirely on x's, and equalize the coefficients of the corresponding monomials.
To show the existence of Q it suffices to remark that any monomial (of F) of degree less than or equal to $2d$ can be written as a product of two elements of z. By writing this in a matriceal form and adding up we obtain a matrix Q. Since a monomial's decomposition into a product of monomials is in general non-unique, Q is not uniquely determined.
To show that Q is affine in λ's and α, note that the system is linear in both the unknowns, $Q_{i,j}$ and α. To be more precise, the linear system can be written as $Av = b - \alpha c$, where $A \in \mathbf{R}^{N_1 \times N_2}$, $b, c \in \mathbf{R}^{N_1 \times 1}$ and $v \in \mathbf{R}^{N_2 \times 1}$ is the vector variable containing the entries of Q, $Q_{i,j}$, $i \leq j$. By solving the system of linear equations in v (using for example Gaussian elimination) we obtain a description of the affine space to which Q belongs.
Note that Q is computed by solving a linear system of $N_1 = \binom{n+2d}{2d}$ equations (this is the number of monomials of degree less than or equal to $2d$ in n variables) with $N_2(N_2+1)/2$ unknowns, where $N_2 = \binom{n+d}{d}$ (the number of monomials of degree less than or equal to d in n variables). □

Let us denote the matrix constructed above $Q(\lambda, \alpha)$, with $\lambda \in \mathbf{R}^k$, $\alpha \in \mathbf{R}$, where $k+1$ is the dimension of the affine space to which Q belongs. As shown, Q is affine in α and λ's.
Let us look at the SDP problem:

$$\begin{array}{ll} \sup & \alpha \\ \text{s.t.} & Q(\alpha, \lambda) \succeq 0. \end{array} \tag{5}$$

Indeed, since $Q(\alpha, \lambda)$ is symmetric, the matrix coefficients of α and λ's will be symmetric matrices. Moreover, $Q(\alpha, \lambda)$ is affine in α and λ's, hence the problem is a standard SDP problem (dual formulation).

The relation between the problems (5) and (4) is studied in the following.

Theorem 4.2 *Let us denote by α_{RAT} the solution of the problem (4), and consequently of the rational optimization problem (3), and by α_{SDP} the solution of (5). Then we have*

$$\alpha_{RAT} \geq \alpha_{SDP}.$$

If $p(x) - \alpha_{RAT} q(x)$ can be written as a sum of squares, then

$$\alpha_{RAT} = \alpha_{SDP}.$$

Proof
Let λ_{SDP} be such that $(\alpha_{SDP}, \lambda_{SDP})$ satisfy (5). Since

$$p(x) - \alpha_{SDP} q(x) = z^T Q(\alpha_{SDP}, \lambda_{SDP}) z \quad \text{and} \quad Q(\alpha_{SDP}, \lambda_{SDP}) \succeq 0$$

we have

$$p(x) - \alpha_{SDP} q(x) \geq 0, \quad \forall x \in \mathbf{R}^n.$$

Hence α_{SDP} satisfies the constraints of (4) and therefore

$$\alpha_{RAT} \geq \alpha_{SDP}.$$

If $p(x) - \alpha_{RAT} q(x)$ can be written as a sum of squares, then there exists a λ_{RAT} such that

$$p(x) - \alpha_{RAT} q(x) = z^T Q(\alpha_{RAT}, \lambda_{RAT}) z, \qquad Q(\alpha_{RAT}, \lambda_{RAT}) \succeq 0.$$

Hence

$$\alpha_{RAT} \leq \alpha_{SDP}.$$

From the result above, equality holds in fact. □

Remark 4.3 *Hilbert showed that there are particular cases in which a positive polynomial can always be written as a sum of squares of polynomials. For homogeneous polynomials, these are:*

$$n \leq 2,\ m = 2 \text{ and } n = 3,\ m = 4$$

where n denotes the number of variables and m the degree of the homogeneous polynomial (see [1], Propositions 6.4.3, 6.4.4). This translates for non-homogeneous polynomials into the following cases

$$n = 1,\ m = 2 \text{ and } n = 2,\ m = 4.$$

Hence, if the polynomial $F(x) = p(x) - \alpha q(x)$ is in one of these cases, we also know that the algorithm will find the infimum, according to Theorem 4.2. If not, then there is always a polynomial $G(x)$ such that $F(x)G^2(x)$ can be written as a sum of squares of polynomials. It is not clear however how to chose the polynomial $G(x)$.

From the practical point of view we are more interested in deciding whether for a particular rational function the infimum was found or just a lower bound of it. The following checking procedure, which makes use of the dual formulation of the SDP problem (5), is indicated in [3]. In fact, one tries to determine a point at which the optimum is attained. If such a point x^* exists then let us denote by z^* the vector z evaluated at x^*. It is not difficult to show that the matrix $X^* = z^{*T} z^*$ is a solution of the dual problem of the SDP. Conversely, from the solution X^* of the dual problem one could determine the vector z^* and further, the point x^* where the optimum is attained. It is argued in [3] that, under no degeneracies, the solution X^* of the dual problem is a matrix of rank 1. In this case, we can determine z^*, for example by performing Gaussian elimination on X^*. Then x^* is found from z^* using the definition of z.

We do not intend here to discuss further such shortcomings of the method. Both its advantages and disadvantages are well explained by their authors (see [4], [3]).

5 Example

Let us consider

$$\inf_{x \in \mathbf{R}^3} \frac{(x_1 + x_2)^4 + {x_1}^3 x_3}{{x_1}^4 + {x_3}^4}.$$

This translates, using (4), into

$$\begin{array}{ll} \sup & \alpha \\ \text{s.t.} & (x_1 + x_2)^4 + {x_1}^3 x_3 - \alpha({x_1}^4 + {x_3}^4) \geq 0, \quad \forall x \in \mathbf{R}^3. \end{array}$$

Since the polynomial is homogeneous of even degree, according to [3] it is sufficient to consider in the vector z, all monomials in the variables x having as degree half the degree of the original polynomial. In our case, this will be 2, hence we define $z = \left[\begin{array}{cccccc} x_1{}^2 & x_2{}^2 & x_3{}^2 & x_1x_3 & x_1x_2 & x_2x_3 \end{array}\right]$.
We compute the symmetric matrix $Q(\lambda, \alpha)$ using the identity of polynomials

$$z^T Q(\lambda, \alpha) z = x_1{}^4 + x_2{}^4 + x_1{}^2 x_2{}^2 - \alpha\,\left(x_1{}^2 + x_3{}^2\right)^2.$$

We obtain

$$Q(\lambda, \alpha) = \left[\begin{array}{cccccc} -\alpha+1 & \lambda_3 & \lambda_4 & 1/2 & 2 & -\lambda_5 \\ \lambda_3 & 1 & \lambda_2 & -\lambda_6 & 2 & 0 \\ \lambda_4 & \lambda_2 & -\alpha & 0 & -\lambda_1 & 0 \\ 1/2 & -\lambda_6 & 0 & -2\lambda_4 & \lambda_5 & \lambda_1 \\ 2 & 2 & -\lambda_1 & \lambda_5 & -2\,\lambda_3+6 & \lambda_6 \\ -\lambda_5 & 0 & 0 & \lambda_1 & \lambda_6 & -2\,\lambda_2 \end{array}\right]$$

With this, problem (5) becomes a standard SDP. Suitable algorithms can be employed for solving it. By running SeDuMi (see [5]) for the above SDP problem, we obtain the solution of (5), -0.5699. Since our problem is one of the special cases mentioned in Remark 4.3, we know that this is the actual infimum.

Let us however, perform the checking procedure as described. We run Gaussian elimination on the solution of the dual problem of (5) and notice that the matrix has indeed rank 1 and

$$z^* = (0.7514, 0.7529, 0.4338, -0.5709, -0.7521, 0.5715).$$

From z^* we recover the solution point $x^* = (-0.8668, 0.8677, 0.6587)$. The rational function evaluated at x^* is equal to the value we have previously found, -0.5699, as expected. We therefore conclude that the infimum of the function is actually attained and one such point is x^*.

6 Conclusions

In this paper we extend an algorithm for global polynomial minimization to the larger class of rational functions. The extension is based on a possibly new result in real algebraic geometry, which allows us to rewrite a rational optimization problem in $\mathbf{R}^n$ as a constrained polynomial optimization problem of a particular type.
Such equivalent formulation of the problem can in principle be solved using a different algorithm than the one discussed here. We have chosen to apply the algorithm of [4], [3] for its possible relevance in applications. The translation of our problem into this setting was immediate.

Acknowledgment I am very grateful to Prof. J. Bochnak for proof-reading Section 2 of the paper.

References

[1] J. Bochnak, M. Coste, M-F. Roy, Géométry algébrique réelle, Springer-Verlag Berlin, 1987.

[2] B. Hanzon, D. Jibetean, Global minimization of a multivariate polynomial using matrix methods, CWI Report PNA-R0109, 2001.

[3] P. Parrilo, Semidefinite programming relaxations for semialgebraic problems, submitted.

[4] N. Shor, Class of global minimum bounds of polynomial functions, Translated from Kibernetica, 6(1987), pp. 9-11.

[5] J. Sturm, SeDuMi 1.03 available at http://fewcal.kub.nl/sturm/software/sedumi.html.