Journal of Global Optimization **18:** 211–218, 2000.

An Algorithm of Global Optimization for Rational Functions with Rational Constraints *

WEIWU FANG[1], TIANJIAO WU[2] and JIANPING CHEN[3]
[1]*DIMACS Center, Rutgers University, Institute of Applied Mathematics, Chinese Academy of Sciences, Beijing, China;* [2]*Institute of Applied Mathematics, Chinese Academy of Sciences, Beijing, China;* [3]*Simon Fraser University, Burnaby, Canada (e-mail: fangww@sina.com.cn)*

(Received 23 November 1999; accepted in revised form 19 May 2000)

Abstract. To solve a system of nonlinear equations, Wu wen-tsun introduced a new formative elimination method. Based on Wu's method and the theory of nonlinear programming, we here propose a global optimization algorithm for nonlinear programming with rational objective function and rational constraints. The algorithm is already programmed and the test results are satisfactory with respect to precision and reliability.

Key words: Nonlinear programming, Global optimization algorithm, Global convergence

1. The problem

Assume that $f(x)$, $g_j(x)$, $x = (x_1, x_2, \dots, x_n)$, $j = 1, 2, \dots, m$, are continuously differentiable, rational and real functions defined on $D \in R^n$. The problem (P) is to find the global minimizers (or global maximizers) of $f(x)$, i.e.,

$$(P): \; st. \quad \begin{array}{l} \min f(x) \\ g_j(x) = 0, \quad j = 1, 2, \dots, l \\ g_j(x) \geq 0 \quad j = l+1, l+2, \dots, m. \end{array}$$

2. On the Wu Elimination

Let $\mathcal{K}[x_1, x_2, \dots, x_n]$ be the ring of polynomials in variables $x_1, x_2, \dots, x_n$ with coefficients in a field $\mathcal{K}$, and this ring is also denoted by $\mathcal{K}[x]$. Consider a fixed ordering on the set of variables: $x_1 \prec x_2 \prec \dots \prec x_n$. For a polynomial $f \in \mathcal{K}[x]$, a variable with the greatest subscript which occurs in *f* is called *a main variable of*

* This work has been supported by the Chinese NSF under the grants NKBRSF G1998030600 and 69772036, and partially supported by DIMACS. DIMACS is a partnership of Rutgers University, Princeton University, AT&T Labs, Bell Labs, NEC Research Institute, and Telcordia Technologies (formerly Bellcore). DIMACS is an NSF Science and Technology Center, funded under contract STC-91–1999; and also receives support from the New Jersey Commission on Science and Technology.

f; further, let *class(f)* be the subscript i of the main variable x_i, and $deg_{x_i}(f)$ be the order of f on the main variable x_i. Given two polynomials $f_1, f_2 \in \mathcal{K}[x]$, f_1 is called *a reduced polynomial with respect to* f_2 if $deg_{x_c}(f_1) < deg_{x_c}(f_2)$, $c = class(f_2)$.

Let PS be a set of polynomials $p_1, p_2, \ldots, p_n$ in the ring $\mathcal{K}[x]$. Denote by $PS = 0$ the system of $p_i = 0, i = 1, 2, \ldots, n$. PS is called *a triangularized polynomial* set if the main variable of the ith polynomial p_i is x_i. By this definition, the polynomial p_i can be written in the form

$$p_i = I_i \cdot x_i^{m_i} + \text{terms of lower degree on } x_i \quad i = 1, 2, \ldots, r,$$

in which the positive integer m_i is the degree of p_i on x_i, the coefficient I_i of the leading term $x_i^{m_i}$ is a polynomial in $\mathcal{K}[x_1, x_2, \cdots, x_{m-1}]$, and is called *the initial of* p_i. A polynomial set AS is called *an ascending set* if for every pair $j > i$, $i = 1, 2, \ldots, r-1$, the degree of initial I_j of p_j on x_i satisfies $deg_{x_i}(I_j) < deg_{x_i}(p_i) = m_i$. A subset BS of an ascending set of PS is called *a basic set of* PS if all main variables of polynomials in the subset BS are different to each other.

For a polynomial p, define *the remainder* $Rem(p/AS)$ *of* p *with respect to a given ascending set* AS by the following division: first, do division of p to each $p_i \subset AS$ with respect to the main variable x_c of p_i, such that the remainder of p is the *reduced polynomial* with respect to p_i; then, the remainder $Rem(p/AS)$ is the set of all remainders of p to each p_i.

For a polynomial set PS, the set of all zeros of PS satisfying $I \neq 0$ is denoted by $Zero(PS/I)$ where I is the product of all I_i and I_i is the initial of p_i. An ascending set CS is called *a characteristic set* of a given polynomial set PS if it satisfies (1) $Zero(PS) \subset Zero(CS)$ and (2) the remainder $Rem(p_i/CS) = 0$, for every polynomial p_i in PS. The following results have been obtained.

LEMMA 1. *(Wu Wen-tsun, 1986). If* $CS = \{C_1, C_2, \ldots, C_r\}$ *is the characteristic set of a polynomial set* PS*, then the zero set of* PS *has the following decomposition*

$$Zero(PS) = Zero(CS/I) + \sum_{i=1}^{r} Zero(PS \cup I_i)$$

where I_i *is the initial of* C_i *and* I *is the product of all* I_i.

LEMMA 2. *(Wu Wen-tsun, 1987, 1992). There is an algorithm which permits to determine characteristic set* CS *for any given polynomial set of* PS *in a finite of steps that satisfies*

$$Zero(CS/I) \subset Zero(PS) \subset Zero(CS)$$

where I *is the product of initials of polynomials in* CS.

Based on the results above, the Wu Elimination for solving a system of equations ($PS = 0$) can be simply described as follows.

(1) Fix an ordering on the set of variables of PS.

(2) Determine the class(p_i) and $deg_{x_c}(p_i)$ of each $p_i \subset PS$.
(3) Select a basic set BS in PS.
(4) Obtain the remainders $Rem((PS - BS)/BS)$.
(5) If $Rem = 0$, then $CS = BS$ and go (6), else let $PS = BS \cup Rem$ and go (2).
(6) Solve $CS = 0$ by the ordinary elimination.

3. The basic steps of algorithm

(1) Transform the original problem into one or several standard problems.

A problem is called *a standard problem* SP if all constraints are equality constraints except slack variables and some independent variables are defined on open sets. The original problem needs to be transformed into one or several of standard problems. The transformation can be done by the following two approaches:

Approach 1. Partition closed sets of variables into open sets. For instance, partition $a_i < x_i \leq b_i$ into $a_i < x_i < b_i$ and $x_i = b_i$.

Approach 2. Add slack variables to transform inequality constraints into equality constraints. For instance, transform $g_j(x) \geq 0$ into $g_j(x) - j_s^2 = 0$ and $-\infty < j_s < \infty$.

By repeatedly using approach 1 and approach 2, the original problem is finally transformed into one or several standard problems SP,

$$\begin{array}{lll} & \min f(x) & \\ st. & h_j(x) = 0, & j \in J \subset \{1, 2, \cdots, m\}, \\ & a_i < x_i < b_i, & i \in K \subset \{1, 2, \cdots, n\}, \\ & -\infty < j_s < \infty, & j_s \text{ are slack variables.} \end{array}$$

(2) Transform each standard problem into two systems of equations.

According to the results about the first-order necessary condition, all local minimizers x^* of a standard problem must be found in the solutions of the system P_1 of equations (corresponding to this standard problem),

$$(P_1) : \begin{cases} \nabla f(x^*) = \sum_{j \in J} \lambda_j \nabla h_j(x^*), \\ h_j(x^*) = 0, \quad j \in J \subset \{1, 2, \ldots, m\}, \end{cases}$$

or P_2,

$$(P_2) : \begin{cases} det|\frac{\partial h_1}{\partial x}, \frac{\partial h_2}{\partial x}, \ldots, \frac{\partial h_p}{\partial x}| = 0 \\ h_j(x^*) = 0 \quad j \in J \subset \{1, 2, \cdots, m\}, \end{cases}$$

where $p = |J|$. Thus, the original problem is transformed into the problem of solving several systems of equations.

(3) Partition the rational expressions of each system of equations into two groups of polynomials.

Partition the rational expression of each system of equations into two polynomial groups **A** and **B**. For instance, assume that $p_i = \frac{k_{ia}(x)}{k_{ib}(x)} = 0, \quad i = 1, 2, \cdots$, then all $k_{ia}(x)'s$ form the group **A** and $k_{ib}(x)'s$ form the group **B**.

(4) Find a characteristic set CS of each group **A** of polynomials.

The procedure of finding the characteristic sets of polynomials can be demonstrated by the following figure:

$$
\begin{array}{llllll}
PS = & PS_0 & & & & \\
 & \left.\begin{array}{l} BS_0(\subset PS_0) \\ RS_0 \end{array}\right] + & = PS_1 & & & \\
 & & \left.\begin{array}{l} BS_1 \\ RS_1 \end{array}\right] + & = PS_2 & & \\
 & & & & \ddots & \\
 & & & & & PS_m \\
 & & & & & BS_m \quad (= CS) \\
 & & & & & RS_m = \Phi
\end{array}
$$

In this figure, $PS_0 = PS$, BS_k is a basic set of PS_k (it is easy to find a basic set according to the definition of basic set), the remainder $RS_k = \{$ the remainder of p to BS_k, $p \in PS_k - BS_k\}$, and $PS_{k+1} = BS_k \cup RS_k$, where $k = 0, \ldots, m-1$.

(5) Find the set of zero points of each characteristic set CS by iterative elimination method.

Since a characteristic set is in triangular form, the zero points of CS can be found by means of iteratively solving a series of algebraic equations with one variable. For an equation with one variable, first using Sturm sequence method, one can determine whether there are real roots of the equation; secondly, the upper and lower bound of real roots can be determined from the theory on polynomials; at last, all real roots of the equation can be found by iterative algorithms, such as bi-part method, accelerated Newton method, and so on.

(6) Exclude some zero points of each CS, which lead to that any $h_j(x) = 0$ does not hold or let any polynomial of the group **B** corresponding to the group **A** be equal to zero. The remainder contains all local minimum points of the problem.

(7) Based on the step (6), find the optimizers by comparing the values of objective function for the zero point sets of all characteristic sets.

4. The global convergence of algorithm

An algorithm is called *globally convergent* if the algorithm can approximate all the global optimizers within any given precision ϵ in finite steps.

THEOREM. *If the problem (P) has the finite number of global optimizers in its feasible region, then the algorithm is globally convergent.*

We give a brief proof as follows.

Proof. 1. By step (1), the original problem can be transformed into the finite number of standard problems in finite steps. By steps (2) and (3), the finite number of polynomial groups can be obtained. Moreover, the finite number of characteristic sets whose order is finite can be obtained in finite steps from Lemma 2.

2. Since the given problem has the finite number of optimizers, each optimizer x^* must be a unique local minimizer in a sufficiently small region $N(x^*)$. That is to say, x^* belongs to the set of isolated zero points of systems of equations P_1 and P_2 (if the gradient vectors $\nabla h_j(x^*)(j = 1, 2, \ldots, p)$ are linearly independent, then the local minimizers are in the set of isolated zero points of P_1; or else they are in the set of isolated zero points of P_2). From $zero(CS/I) \subset zero(PS) \subset zero(CS)$ (Lemma 2), we know that all local minimizers are included in the set of isolated zero points of characteristic sets of P_1 or P_2; i.e., the optimizers belong to the set of isolated zero points of characteristic sets.

3. By the Bezout theorem and the results of algebraic geometry (Shafarevich, 1977), we have known that the set of isolated zero points of a system of polynomial equations is finite, it also means that the set of isolated zero points of a characteristic set is finite. This set of zero points can be obtained by iteratively solving finite number of equations with one variable (the characteristic set is with triangular form): for a polynomial with one variable, we can determine the upper and lower bounds of all real roots by the theory of polynomials, and assuming they are a and b respectively; then, using the bi-part method, we can get a solution with the required precision in less that $log_2 \frac{a-b}{\epsilon}$ steps; in fact, we can use another methods, e.g. Newton method, to find solutions of equations in fewer steps. This means that the algorithm will converge to all zero points of characteristic sets in finite steps.

4. Comparing the values of objective function at n finite solutions can be done in $O(n \cdot \log n)$ steps.

In a word, by steps (2–7) one can find global optimizers in finite steps.

5. The implementation of the algorithm and results

The program of this algorithm involves many mathematical methods and software techniques. We have applied the program to calculate 35 examples in 486/33 microcomputers (4M) and obtained satisfactory results in respect of high precision, reliability, and convenience. Most examples are abstracted from the test examples of twelve books ([1–12]). Each computation result of the thirty-five examples is exactly identical to the original one except for one example, on which there is a little difference at the eighth number after the fraction point from the original result (see Example 3):

$$\begin{array}{lll} the\ original: & x_1 = \pm 0.066041593 & x_2 = \mp 0.192895426 \\ the\ new: & x_1 = \pm 0.066041588 & x_2 = \mp 0.192895426 \\ the\ difference: & \Delta x_1 = \pm 0.00000001 & \Delta x_2 = \mp 0.000000000; \end{array}$$

moreover, the original result is also an approximate result obtained by the interval method. Among these examples, there are different classes of problems, such as, the problems of multi-optimizers, problems of multi-local minimizers, problems of convex or nonconvex programming, problems of geometric programming, problems of general programming with rational functions, and special problems in which the optimizers do not satisfy $K-T$ condition.

We here show five examples of them:

EXAMPLE 1 (Franklin, 1980). Rational functions, two groups of global optimizers.

$$objective = 10 + 6\cdot x_1 + 6\cdot x_1^2 - 22\cdot x_2 - 20\cdot x_1\cdot x_2 + 21\cdot x_2^2 + \frac{1}{25\cdot(1-x_1^2/4-x_2^2)}$$
$$St.\quad x_1 \geq 0.$$

EXAMPLE 2 (Phillips, 1987). Rational functions, four groups of global optimizers.

$$objective = 5\cdot x_1^2 - x_2^2\cdot x_3^4$$
$$St.\quad -2-\frac{5\cdot x_1^2}{x_2^2}+\frac{3\cdot x_3}{x_2}\geq 0$$

EXAMPLE 3 (Hansen, 1992). Two groups of global optimizers.

$$objective = 12\cdot x_1^2 - (63\cdot x_1^4)/10 + x_1^6 + 6\cdot x_1\cdot x_2 + 6\cdot x_2^2$$
$$1-16\cdot x_1^2-25\cdot x_2^2\leq 0,$$
$$-400-145\cdot x_1+13\cdot x_1^3+85\cdot x_2\leq 0,$$
$$-4+x_1\cdot x_2\leq 0$$

EXAMPLE 4 (Van Der Hoek, 1980). A group of global optimizers with multi-local minimizers.

$$\begin{aligned} objective =\ & 14463 + 18340\cdot x_1 + 10197\cdot x_1^2 - 34198\cdot x_2 - 24908\cdot x_1\cdot x_2 + \\ & 20909\cdot x_2^2 + 4542\cdot x_3 - 2026\cdot x_1\cdot x_3 - 3466\cdot x_2\cdot x_3 + 1755\cdot x_3^2 + \\ & 8672\cdot x_4 + 3896\cdot x_1\cdot x_4 - 9828\cdot x_2\cdot x_4 + 2178\cdot x_3\cdot x_4 + 1515\cdot x_4^2 + \\ & 86\cdot x_5 + 658\cdot x_1\cdot x_5 - 372\cdot x_2\cdot x_5 - 348\cdot x_3\cdot x_5 - 44\cdot x_4\cdot x_5 + 27\cdot x_5^2 \end{aligned}$$

$$\begin{aligned} st.\quad & -5+x_1+x_2+x_3+x_4+x_5\leq 5, \\ & -20+10\cdot x_1+10\cdot x_2-3\cdot x_3+5\cdot x_4+4\cdot x_5\geq 0, \\ & -40+8\cdot x_1-x_2+2\cdot x_3+5\cdot x_4-3\cdot x_5\geq 0 \\ & -11+8\cdot x_1-x_2+2\cdot x_3+5\cdot x_4-3\cdot x_5\geq 0, \\ & -30+4\cdot x_1+2\cdot x_2-3\cdot x_3+5\cdot x_4-x_5\leq 0 \end{aligned}$$

EXAMPLE 5 (Schittkowski, 1987). A group of global optimizers with multi-local minimizers.

$$
\begin{aligned}
&objective\ = \\
&\quad 600 + 720 \cdot x_1 + 1260 \cdot x_1^2 - 1072 \cdot x_1^3 - 2454 \cdot x_1^4 + 1344 \cdot x_1^5 + \\
&\quad 952 \cdot x_1^6 - 768 \cdot x_1^7 + 144 \cdot x_1^8 + 720 \cdot x_2 - 4680 \cdot x_1 \cdot x_2 + 7344 \cdot x_1^2 \cdot x_2 + \\
&\quad 5784 \cdot x_1^3 \cdot x_2 - 7680 \cdot x_1^4 \cdot x_2 - 168 \cdot x_1^5 \cdot x_2 + 1344 \cdot x_1^6 \cdot x_2 - 288 \cdot x_1^7 \cdot x_2 + \\
&\quad 3060 \cdot x_2^2 - 19296 \cdot x_1 \cdot x_2^2 + 7776 \cdot x_1^2 \cdot x_2^2 + 9840 \cdot x_1^3 \cdot x_2^2 - \\
&\quad 5370 \cdot x_1^4 \cdot x_2^2 + 2592 \cdot x_1^5 \cdot x_2^2 - 648 \cdot x_1^6 \cdot x_2^2 + 12288 \cdot x_2^3 - \\
&\quad 23616 \cdot x_1 \cdot x_2^3 + 5040 \cdot x_1^2 \cdot x_2^3 + 1240 \cdot x_1^3 \cdot x_2^3 - 4080 \cdot x_1^4 \cdot x_2^3 + \\
&\quad 1224 \cdot x_1^5 \cdot x_2^3 + 14346 \cdot x_2^4 - 11880 \cdot x_1 \cdot x_2^4 + 8730 \cdot x_1^2 \cdot x_2^4 - \\
&\quad 3480 \cdot x_1^3 \cdot x_2^4 + 1305 \cdot x_1^4 \cdot x_2^4 + 1944 \cdot x_2^5 - 1188 \cdot x_1 \cdot x_2^5 + \\
&\quad 3672 \cdot x_1^2 \cdot x_2^5 - 1836 \cdot x_1^3 \cdot x_2^5 - 4428 \cdot x_2^6 + 1944 \cdot x_1 \cdot x_2^6 - \\
&\quad 1458 \cdot x_1^2 \cdot x_2^6 - 648 \cdot x_2^7 + 972 \cdot x_1 \cdot x_2^7 + 729 \cdot x_2^8
\end{aligned}
$$

$$St. \quad -2 \le x_1 \le 2, -2 \le x_2 \le 2$$

6. The features of the program

The program made by us is with the following features:

(1) Global property: for the nonlinear programming problems that both the objective functions and the constraints are rational, global solutions are obtained by a unified and simple way.

(2) Computation precision is high.

(3) Reliability check is easy to be done for each problem by several ways.

(4) A starting iterative point is not required.

(5) Data format is with common form.

(6) Symbol calculation, numerical computation, other mathematical skills, and software techniques are applied in the program.

7. The limitation

The size of polynomials may grow up exponentially while the method of characteristic set is used, so the complexity is exponential in some cases.

8. Acknowledgments

This work has been supported by the Chinese NSF under the grants NKBRSF G1998030600 and 69772036, and partially supported by DIMACS Center. The authors also gratefully thanks Prof. Wu Wen-tsun and Prof. F. Roberts for their help.

References

1. Avriel, M. (1976), *Nonlinear Programming Analysis and Methods*, Prentice-Hall.
2. Franklin, J. (1980), *Methods of Mathematical Economics: linear and nonlinear programming*, Springer, New York.
3. Fletcher, R. (1979), *Practical Method of Optimization*, Wiley and Sons, Chichester.
4. Gabasov, R. and Kirillova, F. (1988), *Methods of Optimization, Optimization Software*, New York.
5. Hansen, E. (1992), *Global Optimization Using Interval Analysis*,, Marcel Dekker, New York.
6. Himmelblan, D.M. (1972), *Applied Nonlinear Programming*, McGraw-Hill, New York.
7. Lperessini, A., Sullivan, F.E. and Uhl, J. Jr. (1992), *The Mathematics of Nonlinear Programming*, World Publishing, New York.
8. Phillips, D. (1987), *Operations Research*, Wiley, New York.
9. Torn, A. (1990), *Global Optimization*, Springer-Verlag, New York.
10. Van Der Hoek, G. (1980), *Reduction Methods in Nonlinear Programming*, Mathematical Center Tracts, Amsterdam.
11. Pardalos, P.M. and Rosen, J.B. (1987), *Constrained Global Optimization: Algorithms and Applications*, Springer-Verlag, Lecture Notes in Computer Sciences 268.
12. Schittkowski, K. (1987), *More Test Examples for Nonlinear Programming Codes*, Springer-Verlag, New York.
13. Shafarevich, I.R. (1977), *Basic Algebraic Geometry*, Springer-Verlag, New York.
14. Wu Tianjiao (1995), *Some Test Problems on Applications of Wu's Method in Nonlinear Programming Problems*, MMRC Preprints, No. 6: P144–155.
15. Wu Wen-tsun (1987), On zeros of Algebraic Equations – an Application of Ritt Principle, *Chinese Scientific Bulletin* 31, 1–5, 1986.
16. Wu Wen-tsun (1987), A Zero Structure Theorem for Polynomial Equations-Solving, *Mathematics and Systems Science* 1: 2–12.
17. Wu Wen-tsun (1992), *On a Finiteness Theorem about Optimization Problems*, MMRC Preprints, No. 8: 1–18.