RT-MAP-7802

EXACT SOLUTION OF SYSTEMS OF LINEAR EQUATIONS WITH ITERATIVE METHODS

Silvio Ursic
Cyro Patarra

FEBRUARY 1978

EXACT SOLUTION OF SYSTEMS OF LINEAR EQUATIONS WITH ITERATIVE METHDS

BY

SILVIO URSIC

and

CYRO PATARRA

An algorithm is presented to compute the exact solution of a system of linear equations with integer coefficients from any method capable of providing a sufficiently accurate approximate solution.

1. INTRODUCTION

Numerical methods for the solution of linear systems of equations are usually classified in two main categories: direct and iterative. Most textbooks on the subject then continue by stating that direct methods are potentially able of finding the exact solution, if exact arithmetic is used, in a finite number of steps. By contrast, iterative methods are presented under the framework that they can only provide us with an approximate solution.

This paper shows that the classification of methods for the solution of linear systems of equations in direct, implying exact, and iterative, implying approximate, is not entirely accurate. In fact we show that any sufficiently close approximation to the solution leads to the exact rational solution of a system with integer coefficients with very little additional work.

As a consequence, the existing iterative methods and their huge supporting literature become available for utilization in the exact solution of systems of linear equations.

2. THE MAIN OBSERVATION

We are interested in finding the exact rational solution to a linear system of equations with integer coefficients. We assume that the system has a unique solution.

The main observation to be made concerns the discrete nature of the problem. The solution vector can be found, for example with Gaussian Elimination, in a finite number of arithmetic operations. As a consequence the numerator and denominator of each rational number in the solution cannot be arbitrarily large. So, there is only a finite number of rationals to be considered as candidates for the solution.

It is therefore possible, in principle, to find the solution to such a system simply by trying one by one all rationals, candidates to the solution. This brute force trial algorithm will obviously have an exponential computing time. Trying one by one all candidates to the solution is not a very good strategy.

The idea is stated more precisely as follows. Let

(1) $$A \cdot x = B$$

be the linear system to be solved. The coefficients of the array A, $a_{i,j}$, $1 \le i, j \le N$, and of the vector B, b_i, $1 \le i \le N$, are integers smaller, in absolute value, than some integer d.

Lemma 1 provides a tight bound of the size of the components of x in (1).

Lemma 1. (Hadamard inequality)

Let det(A) be the determinant of the array A. Then

(2) $$(\det(A))^2 \le \prod_{1\le i\le N}\Big(\sum_{1\le j\le N} a_{i,j}^2\Big).$$

With our bound, d, to the coefficients of (1) we can write

(3) $$|\det(A)| \le N^{N/2} * d^N \le D,$$

for a suitable integer D. For a proof of Lemma 1 see, for example, exercise 4.6.1.15 in (8).

So, if

(4) $$x_i = p_i/q_i, \quad 1 \le i \le N,$$

is the solution to (1), we will have $|q_i| \le D$, $1 \le i \le N$.

Next Lemma tells us that candidates to the solution of (1) are not too close to each other.

Lemma 2. (Minimum distance)

Let p/q and r/s be two rationals with $p/q \neq r/s$ and $|q| \leq D$, $|s| \leq D$. Then

$$(5) \qquad \min \; |p/q - r/s| \leq 1/D^2.$$

Not only there is a finite number of candidates to the solution, but also they are reasonably far apart from each other.

3. A SYSTEM OF INTEGER LINEAR INEQUALITIES AND ITS SOLUTION WITH CONTINUED FRACTIONS.

Let us imagine we were able to find an approximation a/b to the true value p/q of some component of the solution of (1). If the distance between a/b and p/q is less than half the minimum distance between two candidates to the solution, then the nearest candidate to the approximation a/b will be p/q.

More precisely, the system of inequalities

$$(6)\qquad |a/b - \alpha/\beta| \le 1/(2*D^2), \quad 1 \le \beta \le D,$$

with α and β as integer unknowns, has at most one solution. Note that the uniqueness of a possible solution to (6) is guaranteed by Lemma 2.

Inequalities (6) can be rewritten as follows:

$$(7)\qquad \begin{aligned} 2*D^2*b*\alpha - (2*D^2*a + b)*\beta &\le 0, \\ -2*D^2*b*\alpha + (2*D^2*a - b)*\beta &\le 0, \\ \beta &\le D, \\ -\beta &\le -1. \end{aligned}$$

The problem of determining whether a system like (7) has a solution in integers and then, if some solution exists, actually finding one, is in general NP-Complete. Hirschberg and Wong in (4) showed that integer systems of inequalities with only two unknowns are special. They can be solved in polynomial time. In our case it is simpler to find a solution of (7) by going back directly to the continued fraction algorithm.

Continued fractions are an old and venerable topic and have close ties with Euclid's Algorithm. The first documented use of their approximating powers seems to have been done by Huygens, (7). He used continued fractions to compute the best number of teeth in pairs of gears to be used in the driving mechanism of a telescope.

The key result that permits us to solve system (7) efficiently is contained in the following Theorem.

Theorem C. (Continued fractions approximations)

If $|p/q - a/b| \leq 1/(2*q^2)$,

then p/q is a convergent in the continued fraction series for a/b.

Proof: For an algebraic proof see Theorem 184 of (3). For a more geometric and somewhat more revealing approach see Theorem 7.19 in (11).□

Proofs of Theorem C lead directly to Algorithm C.

Algorithm C (Computation of the continued fraction appoximation). Given the integers $a \geq 0$, $b > 0$, $D > 0$, the algorithm computes, when they exist, two integers p', q', such that $|a/b - p'/q'| \leq 1/(2*D^2)$ and $q' \leq D$.

C1. [Initialize.] Set $p \leftarrow 0$, $q \leftarrow 1$, $p' \leftarrow 1$, $q' \leftarrow 0$, $A \leftarrow a$, $B \leftarrow b$.

C2. [Test for end.] If $B = 0$, then go to step C5.

C3. [Compute new approximation.] Set $W \leftarrow \lfloor A/B \rfloor$, $p'' \leftarrow p + W * p'$, $q'' \leftarrow q + W * q'$. If $q'' > D$, then go to step C5.

C4. [Shift and go back.] Set $p \leftarrow p'$, $q \leftarrow q'$, $p' \leftarrow p''$, $q' \leftarrow q''$, $T \leftarrow A - B * W$, $A \leftarrow B$, $B \leftarrow T$; then go back to step C2.

C5. [Test for goodness and terminate.] If $|2 * D^2 * (a * q' - b * p')| \leq b * q'$ then the approximation to a/b is p'/q', otherwise the algorithm returns "NO SOLUTION" and terminates.

Algorithm C is an implementation of the Extended Euclidean Algorithm with the addition of tests in steps C3 and C5.

Step C3 selects one of the continued fraction convergents, namely the one with the largest denominator smaller than the bound D. The choice follows from the following facts:

Fact 1. Each successive convergent, p/q, approximates a/b better and better.

Fact 2. Inequalities (6) have at most one solution.

Hence the only candidate to a solution of (6) is the convergent with the largest possible denominator.

Step C5 is not necessary if we know in advance that two integers, p, q, exist satisfying (6). For then we have

$$|a/b - p/q| \le 1/(2*D^2) \le 1/(2*q^2),$$

because $q \le D$ and we can apply Theorem C.

In general, however, steps C1-C4 may fail to produce the required approximation. The test in step C5 becomes necessary to differentiate between a true solution to (6) and simply a continued fraction approximation p/q to a/b having $q \le D$.

The worst case computing time of Algorithm C is of $O((\log N)^2)$ when all the inputs are bounded by some integer N. For a computing time analysis of Algorithm C, consult (8). For many details of its practical implementation, consult Collins paper (1) and its bibliography.

4. THE PROCEDURE

The results in sections 2 and 3 suggest a two step procedure for the computation of the exact solution of a linear system of equations with integer coefficients.

Step 1. Obtain an approximate solution that differs, in each component, from the true solution by less than $1/(2*D^2)$. D is an integer bound for the denominators of the solution vector.

Step 2. Use Algorithm C to obtain the exact rational solution.

A computing time analysis of step 1 is difficult. The analysis is further complicated by the fact that performance of iterative methods depends strongly on the particular problem being solved. Two improvements of a general nature in the computing time of step 1 are possible.

First, the bound D is in many cases too large. A much smaller bound and considerably less iterations might do. It might be more convenient to apply Algorithm C as a test of termination than to iterate up to the precision necessary to be certain that step 2 will produce the exact solution.

Second, the use of exact arithmetic to compute successive approximations to the solution will cause an increase in the size of the integers to be manipulated during each iteration. Algorithm C can be used to reduce the size of the integers in the approximation. Some care must be exercised, however, not to destroy convergence of the underlying iterative method.

5. CONCLUSION

Continued fractions approximations can be used to obtain the exact rational solution to a problem whenever:

(A) The denominator of the sought rational a/b is bound by some known integer D.

(B) It is possible to obtain an approximation p/q to the rational a/b, satisfying $|p/q - a/b| \le 1/(2*D^2)$.

The continued fractions algorithm closely resembles the Extended Euclidean Algorithm applied to the integers p and q.

As a last observation, consider all the real roots of all the polynomials of degree not greater than N and with integer coefficients bounded by D. Sufficiently small intervals will contain at most one of those roots. We would like to have an efficient algorithm that, given such an interval and given the bounds N and D, would choose a polynomial in our set of polynomials having a root in the given interval. Such an algorithm would allow the use of approximate methods for a wide range of exact computations with algebraic numbers.

The continued fractions algorithm solves the problem efficiently for N = 1.

REFERENCES

1. Collins, G.E. Computer algebra of polynomial and rational functions. Am. Math. Monthly 80, 7 (1973), 725-755.

2. Collins, G.E., and Horowitz, E. The minimum root separation of a polynomial. Math Comp. 28, (1974), 589-597.

3. Hardy, G.E., and Wright, E.M. An introduction to the theory of numbers. Oxford University Press, 1960, 129-151.

4. Hirschberg, D.S., and Wong, C.K. A polynomial-time algorithm for the knapsack problem with two variables. Journal of the ACM 23, 1 (1976), 147-154.

5. Househoulder, A.S. The theory of matrices in numerical analysis. Blaisdell Publ. Co., New York, 1964.

6. Huygens, C. Descriptio automati planetarii. Opuscula Posthuma, Amsterdam, 1728, 174-179.

7. Karp, R.M. On the computational complexity of combinatorial problems. Networks 5, 1 (1975), 45-68.

8. Knuth, D. The art of computer programming. Vol. II: Seminumerical Algorithms. Addison-Wesley, Reading, Mass., 1968.

9. McClellan, M.T. The exact solution of systems of linear equations with polynomial coefficients. Journal of the ACM 20, 4 (1973), 563-588.

10. McClellan, M.T. A comparison of algorithms for the exact solution of systems of linear equations. ACM Transactions on Mathematical Software 3, 2 (1977), 147-158.

11. Stark, H.M. An introduction to number theory. Markham Publ. Co., Chicago, 1971, 181-239.

12. Varga, R.S. Matrix iterative analysis. Prentice Hall, Englewood Cliffs, 1962.

13. Young, D.M. Iterative solution of large linear systems. Academic Press, 1971.

Silvio Ursic
5164 Anton Dr.
Madison, WI U.S.A.

Cyro de Carvalho Patarra
Departamento de Matemática Aplicada
Instituto de Matemática e Estatística
C.P. 20 570 - Ag. Iguatemi
São Paulo - SP

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
Cidade Universitária "Armando de Salles Oliveira"
Caixa Postal n.° 20.570 (Agência Iguatemi) - Tel. 211-0011
SÃO PAULO — BRASIL

"RELATÓRIO TÉCNICO"
DO DEPARTAMENTO DE MATEMÁTICA APLICADA
TÍTULOS PUBLICADOS

RT-MAP-7701 - Ivan de Queiroz Barros
On equivalence and reducibility of Generating Matrices of RK-Procedures - Agosto 1977

RT-MAP-7702 - V.W.Setzer
A Note on a Recursive Top-Down Analizer of N.Wirth - Dezembro 1977

RT-MAP-7703 - Ivan de Queiroz Barros
Introdução a Aproximação Ótima
Dezembro 1977

RT-MAP-7704 - V.W.Setzer, M.M.Sanches
A linguagem "LEAL" para Ensino básico de Computação - Dezembro 1977

RT-MAP-7801 - Ivan de Queiroz Barros
Proof of two Lemmas of interest in connection with discretization of Ordinary Differential Equations
Janeiro 1978

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
Cidade Universitária "Armando de Salles Oliveira"
Caixa Postal n.° 20.570 (Agência Iguatemi) - Tel. 211-0011
SÃO PAULO — BRASIL

RT-MAP-7802 - Silvio Ursic, Cyro Patarra
Exact solution of Systems of Linear Equations with Iterative Methods
Fevereiro - 1978

RT-MAP-7803 - Martin Grötschel, Yoshiko Wakabayashi
Hypohamiltonian Digraphs
Março 1978

RT-MAP-7804 - Martin Grötschel, Yoshiko Wakabayashi
Hypotraceable Digraphs
Maio 1978

RT-MAP-7805 - W.Hesse, V.W.Setzer
The Line-Justifier: an example of program development by transformations
Junho 1978

RT-MAP-7806 - Ivan de Queiroz Barros
Discretização
Cap.I - Tópicos Introdutórios
Cap.II - Discretização
Julho 1978

RT-MAP-7807 - Ivan de Queiroz Barros
(Γ', Γ)-Estabilidade e Métodos Preditores-Corretores
Setembro 1978

RT-MAP-7808 - Ivan de Queiroz Barros
Discretização
Cap. III - Métodos de passo progressivo para Eq. Dif. Ord. com condições iniciais
Setembro 1978

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
Cidade Universitária "Armando de Salles Oliveira"
Caixa Postal n.º 20.570 (Agência Iguatemi) - Tel. 211-0011
SÃO PAULO — BRASIL

RT-MAP-7809 - V.W.Setzer
Program development by transformations applied to relational Data-Base queries
Novembro 1978

RT-MAP-7810 - Nguiffo B. Boyom, Paulo Boulos
Homogeneity of Cartan-Killing spheres and singularities of vector fields
Novembro 1978

RT-MAP-7811 - D.T.Fernandes e C. Patarra
Sistemas Lineares Esparsos, um Método Exato de Solução
Novembro 1978

RT-MAP-7812 - V.W.Setzer e G.Bressan
Desenvolvimento de Programas por Transformações: uma Comparação entre dois Métodos
Novembro 1978

RT-MAP-7813 - Ivan de Queiroz Barros
Variação do Passo na Discretização de Eq.Dif. Ord. com Condições Iniciais
Novembro 1978

RT-MAP-7814 - Martin Grötschel e Yoshiko Wakabayashi
On the Complexity of the Monotone Asymmetric Travelling Salesman Polytope I: HIPOHAMILTONIAN FACETS
Dezembro 1978

UNIVERSIDADE DE SÃO PAULO

INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

Cidade Universitária "Armando de Salles Oliveira"
Caixa Postal n.° 20.570 (Agência Iguatemi) - Tel. 211-0011
SÃO PAULO — BRASIL

RT-MAP-7815 - Ana F. Humes e E. I. Jury

Stability of Multidimensional Discrete Systems: State-Space Representation Approach

Dezembro 1978