bounds and algorithms

Loïc Pottier

January 9, 1991

S.A.F.I.R. Project
Institut National de Recherche en Informatique et Automatique
2004 route des Lucioles, Sophia Antipolis, 06565 Valbonne CEDEX, FRANCE
Email : pottier@mirsa.inria.fr

Abstract : We give new bounds and algorithms for minimal solutions of linear diophantine systems. These bounds are simply exponential, while previous known bounds were, at least until recently, doubly exponential.

1 Introduction

A linear diophantine system $Ax \leq b$ is a set of inequations with integer coefficients (A is a matrix of integers with m rows and n columns, b a vector of $Z^m$and x a vector of n indeterminates), whose we search integer solutions.
Recall that to decide if such a system has at least one integer solution is NP-complete (it is the NP-completeness of integer linear programming).
We are interested here in describing and computing the set of solutions. Remark that these systems arise in pattern matching compilation theory.
We will reduce our problem to the study of an equivalent problem, which is solving in non-negative integers the systems $Ax = 0$ (Frobenius problem). These systems arise in several sub-fields of equational rewriting theory, for instance in associative-commutative unification, or in Makanin algorithm. Remark that solving systems $Ax = b, x \geq 0$, occurring in AC-unification, reduces also to Frobenius problem by adding a new variable to x.
The non-negative integer solutions of $Ax = 0$ form a sub-monoid M of N^n, generated by its non zero minimal elements for the partial order $(x_1, \ldots, x_n) \preceq (y_1, \ldots, y_n) \iff \forall i, 1 \leq i \leq n, x_i \leq y_i$. They form a finite set. We will call this set "the Hilbert basis of M" (after [F.Giles and W.R.Pulleyblank 79]), and denote it by $\mathcal{H}(M)$.
In the two next sections we will bound and compute the elements of $\mathcal{H}(M)$. The last section applies the previous results to the initial problem, i.e. the resolution of systems $Ax \leq b$.

simply exponential in the size of A (for example $n.m.(log\|A\|_\infty + 2)$).
But we are interested here to uniformly bound the norms of the elements of $\mathcal{H}(M)$.
Let

$$\|M\|_\infty = sup_{x\in\mathcal{H}(M)}\|x\|_\infty$$

and

$$\|M\|_1 = sup_{x\in\mathcal{H}(M)}\|x\|_1$$

(with $\|x\|_1 = \sum_i \mid x_i \mid$).
[I.Borosh and L.B.Treybig 76] have upper bounded $\|M\|_\infty$ with an expression which is doubly exponential in the size of A.
As far as we know, two simply exponential bounds exist to bound $\|M\|_\infty$ or $\|M\|_1$.
We have given the first one in [L.Pottier 90].
The second one can be deduced from an rather unknown result of [J.L.Lambert 87], and has been found independently by [E.Domenjoud 90] in a better form.
These two bounds are essentially different, because of their expressions and their proofs.
We give here two new finest bounds, the first one inspired from [L.Pottier 90], the second one from [J.L.Lambert 87] and [A.Koscielski and L.Pacholski 90],.
Recall that in the case of one equation ($m = 1$), [G.Huet 78] and [J.L.Lambert 87] have given bounds only depending of $\|A\|_\infty$. In the case of two equations, [J.F.Romeuf 89] gave a bound which is quadratic in the size of A.

2.1 First bound

This bound is inspirated by [L.Pottier 90].
Let $\|A\|_{1,\infty} = sup_i\{\sum_j \mid a_{ij} \mid\}$, and let r be the rank of A.

Theorem 1

$$\|M\|_1 \leq (1 + \|A\|_{1,\infty})^r = B_0$$

Proof :

We can without restriction choose r independent equations of $Ax = 0$. Let $x = (x_1, \ldots, x_n)$ be a non zero element of M, $p = \|x\|_1$, and $\{e_1, \ldots, e_n\}$ be the canonical basis of R^n. For every y in R^n, we note C_y the cube of volume 1 defined by

$$z \in C_y \Leftrightarrow z = y + \sum_{i=1}^{n} \lambda_i e_i,\ \forall i \in [1, n],\ \lambda_i \in [0, 1]$$

We will recursively define a sequence $y^0, \ldots, y^p$ of N^n and a sequence $z^0, \ldots, z^p$ of R^n verifying :

$y^0 = z^0 = 0$ are clearly convenient .

Suppose we have built y^k, and $0 \leq k \leq p-1$. $[0,x] \cap C_{y^k}$ is the set of all z which writes $\lambda \sum_i x_i e_i$ with $0 \leq \lambda \leq 1$ and

$$\forall i \in [1,n],\ y_i^k \leq \lambda x_i \leq y_i^k + 1$$

It is by hypothesis non empty, it is a segment. Take $z^{k+1} = \lambda_k \sum_i x_i e_i$ its bound where λ_k is maximum. x is non zero, then there exists a j such that

$$\lambda_k = \frac{y_j^k + 1}{x_j} = inf_{i|x_i \neq 0}\{\frac{y_i^k + 1}{x_i}\}$$

Now, let $y^{k+1} = y^k + ej$. We have now

$$\forall i \in [1,n],\ y_i^{k+1} \leq \lambda_k x_i \leq y_i^{k+1} + 1$$

and z^{k+1} belongs then to the cube $C_{y^{k+1}}$.

The points z^{k+1} and y^{k+1} are then correctly built. Finally, if $k = p-1$, then $y^p = x$, because $y^p \preceq x$ and $\|y^p\|_1 = p = \|x\|_1$, and we take $z^p = x$.

Now, let $y'^k = z^k - y^k$. We have now :

$$\forall i, 0 \leq y_i'^k \leq 1$$

then, if $(Ay^k)_i$ is the ith coordinate of Ay^k :

$$|\ (Ay^k)_i\ | = |\ (Az^k)_i - (Ay'^k)_i\ | = |\ (Ay'^k)_i\ |$$

As $0 \leq y_i'^k \leq 1$, there is then at most $\sum_j |\ a_{ij}\ | + 1$ possible values for $(Ay^k)_i$ and then at most B_0 distinct vectors Ay^k.

Now, suppose $p > B_0$. By the pigeon holes principle, it exists then i and j, $i > j > 0$ with $Ay^i = Ay^j$. Let $z = y^i - y^j$. We have now $Az = 0$. More we have $0 \prec z \prec x$, and $z \in M$. Then $x \notin \mathcal{H}(M)$.

□

2.2 Second bound

Let a_{ij} be the term of row i and of column j of the matrix A, and $\|A\|_1 = \sum_{i,j} |\ a_{ij}\ |$. Let D be the largest absolute value of the minors of A. [J.L.Lambert 87] gives the following result :

Theorem 2 (Lambert)

$$\|M\|_\infty \leq nD$$

$$\|M\|_\infty \leq (n-r)D_r = B_1$$

and then

$$\|M\|_\infty \leq (n-r)\left(\frac{\|A\|_1}{r}\right)^r = B_2$$

Remark : the first bound is the same as the bound of [E.Domenjoud 90], found independently.

Proof :

Let $\mathcal{C}$ be the cone of R^n of non-negative real solutions of $Ax = 0$. Let $\mathcal{C}_j$ be its intersection with the hyperplane of equation $x_j = 0$. It is clear that $\mathcal{C}$ is the convex hull of the union of the cones $\mathcal{C}_j$. We can recursively apply this decomposition of $\mathcal{C}$ to the $\mathcal{C}_j$, while the dimension of built cones is largest than 1. $\mathcal{C}$ is then the convex hull of the union of these cones of dimension 1, called "edges" of $\mathcal{C}$.

Every of these edges is then the set of non-negative solutions of a system of equations obtained by choosing r independent equations of $Ax = 0$, and by adding to them $n-r-1$ equations of type $x_j = 0$ in order to keep the system of maximum rank, i.e. $n-1$.

We can then obtain director vectors (with non-negative integer coefficients) of edges by computing the n minors of order $n-1$ for every the preceding systems, that reduces to compute minors of order r of A.

Let $g_1, \ldots, g_k$ be these vectors, which have then their coordinates upper bounded in absolute value by D_r.

M is then included in the non-negative cone that they generate (the linear combinations with non-negative real coefficients), which is exactly $\mathcal{C}$, with dimension at most $n-r$. We have then, with the theorem of Carathéodory :

$$Ax = 0, x \geq 0 \implies \exists j_1, \ldots, j_{n-r}, \exists \alpha_1, \ldots, \alpha_{n-r} \geq 0, x = \sum_{l=1}^{n-r} \alpha_i g_{j_i}$$

If now x is minimal, it is clear that the α_i are strictely smaller than 1. We obtain then the first part of the theorem.

The second part is a simple upper bound of the determinant of a square submatrix A' of order r of A :

$$\mid det(A') \mid \leq \prod_j \sum_i \mid a'_{ij} \mid \leq \left(\frac{\sum_{i,j} \mid a'_{ij} \mid}{r}\right)^r \leq \left(\frac{\|A\|_1}{r}\right)^r$$

The bound B_1 can be optimal , as we will see on examples, but it is not reasonably computable in practice: is it better to compute *all* the principal minors of A than to directly compute $\mathcal{H}(M)$, for example with the algorithm of [E.Contejean and H.Devie 89] which does not use a bound of $\mathcal{H}(M)$?

2.3 Comparison of B_0, B_1 and B_2

It is clear that these three bounds are simply exponential in the size of A. The following examples show that we can not compare in general the first and the last, the second being sometimes optimal, but being not computable in practice.
We have the following inequalities :

$$\|x\|_1 \leq B_0$$

$$\|x\|_1 \leq nB_2$$

$$\|x\|_\infty \leq B_0$$

$$\|x\|_\infty \leq B_2$$

So we will study the behaviours of the ratios $\frac{B_2}{B_0}$ (bounds of $\|x\|_\infty$) and $\frac{nB_2}{B_0}$ (bounds of $\|x\|_1$), when n or $\|A\|_\infty$ increases to infinity.

2.3.1 Example 1

Let a be an integer greater or equal to 3 and A the matrix

$$\begin{pmatrix} a & 1-a & & \\ & \ddots & \ddots & \\ & & a & 1-a \end{pmatrix}$$

where the non written coefficients are zero.
We have $r = m = n-1$ and $\mathcal{H}(M)$ has only one element : $((a-1)^{n-1}, a(a-1)^{n-2}, \ldots, a^{n-1})$.
Then :

$$\|M\|_\infty = B_1 = a^{n-1}, \|M\|_1 = a^n - (a-1)^n, B_2 = (2a-1)^{n-1}, B_0 = (2a)^{n-1}$$

B_1 is then optimal, B_2 and B_0 being very close.
Asymptoticaly, we have finally :

$$\lim_{n\to\infty} \frac{B_2}{B_0} = 0, \lim_{n\to\infty} \frac{nB_2}{B_0} = 0, \lim_{a\to\infty} \frac{B_2}{B_0} = 1, \lim_{a\to\infty} \frac{nB_2}{B_0} = n$$

$$A = \begin{pmatrix} 1 & \dots & 1 & & & & & -1 \\ & & & 1 & \dots & 1 & & -1 \\ & & & & & & \ddots & \vdots \\ 1 & & & 1 & & & & -1 \\ & \ddots & & & \ddots & & \dots & \vdots \\ & & 1 & & & 1 & & -1 \end{pmatrix}$$

where the non written coefficients are zero.
We have $r = 2k - 1$. The Hilbert basis of M is the set of matrices of permutations of order k (cf [R.P.Stanley 83]). Then :

$$\|M\|_\infty = 1, \|M\|_1 = k+1, B_1 \geq k^2 - 2k + 2$$

$$B_2 = (k^2 - 2k + 2)\left(\frac{2k(k+1)}{2k-1}\right)^{2k-1}, B_0 = (k+2)^{2k-1}$$

and :

$$\lim_{n\to\infty} \frac{B_2}{B_0} = \infty, \lim_{n\to\infty} \frac{nB_2}{B_0} = \infty$$

which gives the inverse behaviour of the preceding example.

3 Algorithms

The subject of this section is the computation of all the elements of $\mathcal{H}(M)$.
The first algorithms are based on the bounds of [G.Huet 78] and [J.L.Lambert 87] relative to one equation, extended to a system of equations, but giving then doubly exponential bounds. They are the followings :

Property 1 *Let $x = (x_1, \dots, x_p, y_1, \dots, y_q)$ be an element of the Hilbert basis of the equation*

$$a_1x_1 + \dots + a_px_p + b_1y_1 + \dots + b_qy_q = 0.$$

where the a_i are non-negative and the b_j are negative. Then :

$$\forall i, \mid x_i \mid \leq sup_j \mid b_j \mid$$

(Huet)

$$\sum_i x_i \leq sup_j \mid b_j \mid$$

(Lambert).
(the part concerning the y_j is symmetric).

automaton enumerating $\mathcal{H}(M)$, and a quadratic bound in this case.

3.1 Algorithm of Contejean-Devie

[E.Contejean and H.Devie 89] have found a elegant algorithm which does not need any bound of $\mathcal{H}(M)$. The principle is the following. Let us order N^n by the order $\preceq$ defined before, and obtain a DAG (directed acyclic graph) of root 0. The algorithm enumerates a part of this DAG with the following principle :
begin with 0, and if the current vertex is a non zero vector x such that for no one among its ancestors y we have $A(x-y)=0$, visite its sons $x+e_j$ verifying $Ax.Ae_j \leq 0$ (the . denoting the scalar product of R^n).
This algorithm suprizingly terminates and is complete. If we do not visite twice a vertex of the DAG, and keep only minimal solutions for $\preceq$, we then obtain $\mathcal{H}(M)$.
Different refined versions of this algorithm exist, which eliminate early in the process some unusefull parts of the DAG.
The only result of complexity about this algorithm is, to our knowledge, a consequence of [L.Baratchart and L.Pottier 89], which gives a doubly exponential bound on the number of visited vertices.
This algorithm has good behaviour in practice, but is expensive if the elements of $\mathcal{H}(M)$ have large norms.

3.2 Algorithm of Domenjoud

In [E.Domenjoud 90] is described an algorithm which only builds solutions of $Ax=0$ to compute minimal solutions (as the second algorithm that we present does). This recent algorithm would be interesting in pratice.

3.3 An algorithm inspirated by theorem 3

The analysis of the proof of the theorem 1 allows to modify the method of the algorithm of [E.Contejean and H.Devie 89] in only increment x by the e_l such that for every i, the i-th coordinate of $A(x+e_l)$ is between $-\sum_j a_{ij}^+$ and $\sum_j a_{ij}^-$.
The generators are then all obtained as points of the sequences strictly increasing built similarly to the preceding algorithm.

3.4 Use of standard basis

We give here a new algorithm using the preceding bounds on $||M||_\infty$ and $||M||_1$, based on the theory of standard basis (or Gröbner basis).
Let us recall basic notions of standard basis.
For a polynomial P of the ring $K[X_1,\ldots,X_n]$, we note $in(f)$ the maximum monomial of f w.r.t a choosen admissible ordering on monomials (i.e. a total ordering stable by

[A.Galligo 85]).
In our problem, the idea is to see the columns of A as the exponents of monomials in m variables, and the solutions of $Ax = 0$ in Z^nas sysygies relative to these monomials. This idea has been introduced by [F.Ollivier 90] for computation of standard basis of sub-algebras. Then a computation of an appropriate standard basis gives a canonical rewriting system whose the inverse enumerates M by increasing norm. Finaly it suffices to only keep the minimal solutions for $\preceq$ and of norm smaller than $inf\{nB_2, B_0\}$.
Let $T, X_1, \ldots, X_m, Y_1, \ldots, Y_n$ be $n+m+1$ variables, and k be an arbitrary field.
We note a_j for the jth column of A.
For all $\alpha \in Z^m$ and $\beta \in Z^n$, we note X^α and Y^β the monomials $X_1^{\alpha_1} \ldots X_m^{\alpha_m}$ and $Y_1^{\beta_1} \ldots Y_n^{\beta_n}$.
α^+ is the maximum of α and zero (for the partial order $\preceq$), and α^- is the maximum of $-\alpha$ and zero. Then $\alpha = \alpha^+ - \alpha^-$.
For every $j \in [1, n]$, we define a polynomial P_j in the ring $R = k[T, X_1, \ldots, X_m, Y_1, \ldots, Y_n]$:

$$P_j = X^{a_j^+} - Y_j X^{a_j^-}$$

Let $\mathcal{I}$ be the ideal of R generated by the P_j and the polynomial $P_0 = TY_1 \ldots Y_n - 1$, and $\mathcal{J}$ its trace (i.e. its intersection) on the ring $R' = k[Y_1, \ldots, Y_n]$.
Now, let $\mathcal{B}_\mathcal{I}$ be the reduced standard basis of $\mathcal{I}$ for the following ordering on the monomials of R :
we compare first lexicographically the X_i, and in case of equality we use the degree order, and finally the lexicographic order.
Let $\mathcal{B}_\mathcal{J}$ be the set of polynomials of $\mathcal{B}_\mathcal{I}$ where the X_i's and T do not appear.
$\mathcal{B}_\mathcal{J}$ is then a standard basis of the ideal $\mathcal{J}$ for the degree order (from a remark of D.Bayer and M.Stillman). More, its elements are differences of monomials (because those of $\mathcal{B}_\mathcal{I}$ are).
Then let $Y^{\alpha_k} - Y^{\beta_k}$ be the elements of $\mathcal{B}_\mathcal{J}$, $k \in [1, p]$ and Y^{α_k} being the leading monomials.
Now, note $\longrightarrow$ the rewriting relation corresponding to the division of polynomials by the standard basis $\mathcal{B}_\mathcal{J}$, and $\xrightarrow{*}$ its transitive reflexive closure.
We write $m1 \downarrow m2$ when two monomials $m1$ and $m2$ rewrite in the same monomial, or equivalently when $m1 - m2 \xrightarrow{*} 0$.
Then :

Property 2

$$\forall x \in Z^n \ , \ Ax = 0 \iff Y^{x^+} - Y^{x^-} \in \mathcal{I} \iff Y^{x^+} \downarrow Y^{x^-}$$

Proof :

ing system equivalent to the relation $P = Q \Leftrightarrow P - Q \in \mathcal{J}$. $\square$

As a consequence :

Property 3

$$\forall x \in N^n \ , \ x \in M \iff Y^x \xrightarrow{*} 1$$

This last property allows to test if M is non reduced to $\{0\}$:

Theorem 4 *The system $Ax = 0$ has a positive solution if and only if it exists in $\mathcal{B}_{\mathcal{I}}$ a polynomial of the form $Y^\alpha - 1$.*

More, we have an effective representation of M with of rewriting rules :
Let SR_M the system of rewriting rules on monomials obtained in reversing the polynomials of $\mathcal{B}_{\mathcal{J}}$:

$$SR_M = \{Y^{\beta_1} \longrightarrow Y^{\alpha_1}, \ldots, Y^{\beta_p} \longrightarrow Y^{\alpha_p}\}$$

Note $\longrightarrow_i$ its rewriting relation (it is the symmetric of $\longrightarrow$, and it is not noetherian).
Then

$$x \in M \iff 1 \xrightarrow{*}_i Y^x$$

We can then generate all the elements of M by exploration of the tree of rewritings of 1 by $\longrightarrow_i$, and obtain $\mathcal{H}(M)$ in only keeping the minimal elements of degree smaller than the bounds nB_2 and B_0 (This method is complete because $\longrightarrow_i$ increases the degrees of monomials, and then the norms $\|.\|_1$ of the solutions).
More precisely :

Theorem 5 *The following algorithm stops and returns $\mathcal{H}(M)$:*

1. $E := \{1\}$
2. `While` $\exists x \in E, y \notin E, with\ x \longrightarrow_i y,\ and\ deg(y) \leq inf\{nB_2, B_0\}$

   `Do` $E := E \cup \{y\}$
3. `Return` $\mathcal{H}(M)$ $:=$ *minimal elements for $\preceq$ of vectors of exponents of monomials of $E - \{1\}$.*

Corollary 1 *It exists two finite parts $\mathcal{C}_1$ and $\mathcal{C}_2$ of Z^n such that :*

$$x \in \mathcal{C} \Leftrightarrow x = x_1 + x_2 + \ldots + x_k, \text{ with } x_1 \in \mathcal{C}_1, \text{ and } x_2, \ldots, x_k \in \mathcal{C}_2$$

and

$$\forall x \in \mathcal{C}_1 \cup \mathcal{C}_2, \ \|x\|_1 \leq (2 + \|A\|_{1,\infty} + \|b\|_\infty)^m$$

Proof :

We will reduce the problem to solve in N a system of homogeneous equations.

Let ψ be an endomorphism of R^n which only change the signs of some coordinates of its argument, and $\psi(A)$ the obtained matrix when changing the signs of the corresponding columns of A.

Let $y = (y_1, \ldots, y_m)$ be a vector of m new variables, z a last variable, t the vector obtained in catenating x, y,and z, and let ϕ be the projection mapping t in x.

Let A' be the matrix obtained in catenating $\psi(A)$, the identity of order m and the opposite of b.

We have now clearly the equivalence :

$$Ax \leq b \ \psi(x) \in N^n \iff \exists t \in N^{n+m+1}, A't = 0, z = 1, x = \psi(\phi(t))$$

More $rank(A') = m$, and $\|A'\|_{1,\infty} \leq \|A\|_{1,\infty} + 1 + \|b\|_\infty$.

Let $\mathcal{H}$ the Hilbert basis of $A't = 0$, and $\mathcal{C}_1^\psi$ (resp. $\mathcal{C}_2^\psi$) the image by ϕ of the elements of $\mathcal{H}$ such that $z = 1$ (resp. $z = 0$).

We take then $\mathcal{C}_1$ (resp. $\mathcal{C}_2$) equal to the union of the $\mathcal{C}_1^\psi$ (resp. $\mathcal{C}_2^\psi$) for the 2^n possible choices of ψ.

As $\|\psi(x)\|_1 = \|x\|_1$ and $\|\phi(t)\|_1 \leq \|t\|_1$, we obtain the second part of the result. $\square$

5 Acknowledgements

We would like to thank C.Traverso and A.Galligo for usefull discussions about first versions of algorithm 3.3., J.F.Romeuf who indicated us theorem 2, and the referees for their remarks on the first version of this paper.

[I.Borosh and L.B.Treybig 76] "Bounds of non-negative integral solutions of linear diophantine equations", Proc. AMS v.55, n.2, march 1976.

[B.Buchberger 83] "Gröbner basis: an algorithmic method in polynomial ideal theory" Camp. Publ. Nr. 83-29. 0, nov. 1983.

[E.Contejean and H.Devie 89] "Solving systems of linear diophantine equations", UNIF'89, proc. of the third international Workshop on unification, Lambrecht, RFA 89.

[E.Domenjoud 90] "Solving Systems fo Linear Diophantine Equations : An Algebraic Approch", UNIF'90, International Workshop on Unification, Leeds UK, july 1990.

[A.Galligo 85] "Algorithmes de calcul de bases standard", Preprint Université de Nice, France, 1985.

[F.Giles and W.R.Pulleyblank 79] "Total dual integrality and integer polyedra" Linear algebra and its applications, 25, pp191-196, 1979.

[G.Huet 78] "An algorithm to generate the basis of solutions to homogeneous linear diophantine equations", Information Processing Letters, vol.3, No.7, 1978.

[A.Koscielski and L.Pacholski 90] "Exponent of periodicity of minimal solutions of word equations", manuscript, University of Wroclaw, Poland, june 1990.

[J.L.Lambert 87] "Une borne pour les générateurs des solutions entières positives d'une équation diophantienne linéaire." Comptes Rendus de l'Académie des Sciences de Paris, t.305, Série I, pp39-40, 1987.

[J.L.Lambert 87] "Un problème d'accessibilité dans les réseaux de Petri" Phd thesis, theorem I.5., p 18, University of Paris-Sud, Orsay, France, 1987.

[F.Ollivier 90] "Le problème de l'identifiabilité structurelle globale : approche théorique, méthodes effectives et bornes de complexité", Phd Thesis, Ecole Polytechnique, France, june 1990.

[L.Pottier 90] "Bornes et algorithmes de calcul des générateurs des solutions de systèmes diophantiens linéaires", internal report, INRIA, feb. 90, Comptes Rendus de l'Académie des Sciences de Paris, t.311, Série I, p813-816,1990.

[J.F.Romeuf 89] "Solutions of a linear diophantine system", UNIF'89, proc. of the third international Workshop on unification, Lambrecht, RFA 89.

[R.P.Stanley 83] "Combinatorics and commutative algebra", Progress in Mathematics, Birkäuser ed., 1983.