On systems of several equations modulo three

A. V. Seliverstov, O. A. Zverkov
Institute for Information Transmission Problems
of the Russian Academy of Sciences
(Kharkevich Institute)

XIV Belarusian Mathematical Conference
Minsk
October 31, 2024

Let us denote by $GF(3)$ the field of residues modulo three.
Elements of the field $GF(3)$ are numbers $\{0, 1, 2\}$.
For example, let us write $-1 = 2$ instead of $-1 \equiv 2 \pmod 3$.
A solution to a system of equations in which the value of each variable belongs to the set $\{0, 1\}$ is called a $(0, 1)$-solution.
The recognition problem of deciding whether there is a $(0, 1)$-solution to a system of linear equations over the field $GF(3)$ is NP-complete.
However, for a single equation, this problem can be easily solved: only a linear equation of the type $x_k = 2$ does not have a $(0, 1)$-solution because each linear equation that depends non-trivially on two or more variables has a $(0, 1)$-solution.

x	y	$x + y$	$x + 2y$
0	0	0	0
0	1	1	2
1	0	1	1
1	1	2	0

Proposition. *There is an algorithm that takes as input a system of m linear equations in n variables over the field $GF(3)$ and in $n^{O(m)}$ time accepts the input if and only if the system has a $(0,1)$-solution.*

Any $(0,1)$-solution to a system in n variables over $GF(3)$ extends to a $(0,1)$-solution to a system in $n + m\lceil \log_2(n+1) \rceil$ variables over $\mathbb{Z}$. Each system has m equations. Here the j-th equation

$$a_{j0} + a_{j1}x_1 + \cdots + a_{jn}x_n = 0$$

over $GF(3)$ corresponds to the equation

$$a_{j0} + a_{j1}x_1 + \cdots + a_{jn}x_n = 3\left(y_{j1} + 2y_{j2} + \cdots + 2^k y_{jk}\right)$$

over $\mathbb{Z}$, where $k = \lceil \log_2(n+1) \rceil$ and each new variable $y_{j1}, \ldots, y_{jk}$ occurs only once. The absolute values of the coefficients are $O(n)$. The search for a $(0,1)$-solution can be performed using dynamic programming in at most $n^{O(m)}$ time.

The proposition is proven.

Let a system of linear equations in variables $x_1, \ldots, x_n$ contain more than one equation and some equation non-trivially depends on x_k.

Definition. A new system of linear equations is obtained from the original system by eliminating the variable x_k if two conditions hold:
(1) the new system does not depend on the variable x_k and
(2) the original system is equivalent to the union of the new system and exactly one equation (depending on x_k) equal to a linear combination of the equations of the original system.

$$\begin{cases} x_1 + x_2 & = & 1 \\ x_1 - x_2 + x_3 + x_4 & = & 0 \end{cases}$$

Eliminating the variable x_3 yields one equation:

$$x_1 + x_2 = 1$$

and each of its $(0,1)$-solutions can be extended to a $(0,1)$-solution to a system of two equations.

Eliminating a variable may result in a system having a larger number of $(0,1)$-solutions than the original system had.

Lemma 1. *Given natural numbers n and m satisfying inequalities $n \geqslant 5$, $m \geqslant 2$, and $m \leqslant \log_3(2n-1)$, and a system of m linear equations in n variables over the field $GF(3)$. Suppose for each index $1 \leqslant k \leqslant n$, there exists an equation that depends non-trivially on the variable x_k. If this system has no $(0,1)$-solution, then there exists an index $k \leqslant n$ such that eliminating the variable x_k yields a new system that has no $(0,1)$-solution. Moreover, this new system can be found in polynomial time $O(mn\lceil\log_2(n+1)\rceil)$.*

For a system $A\mathbf{x} = \mathbf{b}$, the variables corresponding to the columns in the matrix A that are proportional to each other are eliminated.

It is possible over the field $GF(3)$.

The proof of Lemma 1 uses the following claim.

Claim. *Given an $m \times n$ matrix A without zero columns.*
If $m \leqslant \log_3(2n-1)$, then there are two linearly dependent columns.
These columns can be found in polynomial time.

The number of possible different non-zero columns is $3^m - 1$.
This set is divided into $(3^m - 1)/2$ pairs of linearly dependent columns.
Therefore, the fulfillment of the condition $n \geqslant (3^m + 1)/2$ ensures that there are two linearly dependent columns in the matrix A.
Let us denote by j and k the numbers of these columns. One can find the numbers j and k by going through $n(n-1)/2$ variants and checking the linear dependence of the corresponding columns.
But another approach is faster.
In almost linear $O(mn \log_2 n)$ time, all columns are sorted.
Then, for the j-th column, the search for the k-th column is performed in $O(m \log_2 n)$ time.

The original system $A\mathbf{x} = \mathbf{b}$ is equivalent to the system $B\mathbf{x} = \mathbf{c}$, where in the $m \times n$ matrix B in columns with numbers j and k nonzero entries are located only in one row with number ℓ. For example, for $j = 1$, $k = 2$ and $\ell = 1$:

$$B = \begin{pmatrix} b_{11} & b_{12} & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & * & \cdots & * \end{pmatrix},$$

where $b_{11} \neq 0$ and $b_{12} \neq 0$.

If a new system obtained by removing the ℓ-th equation from this system has a $(0, 1)$-solution, then it has a $(0, 1)$-solution for any values of the variables x_j and x_k. Consequently, the entire system also has a $(0, 1)$-solution, since the choice of values of the variables x_j and x_k allows the ℓ-th equation to be satisfied for any evaluation of the remaining variables.

Removing the ℓ-th equation corresponds to eliminating the variables x_j and x_k.

Lemma 1 is proven.

Lemma 2. *Let us consider a system of linear equations in n variables over the field $GF(3)$:*

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ \cdots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{cases}$$

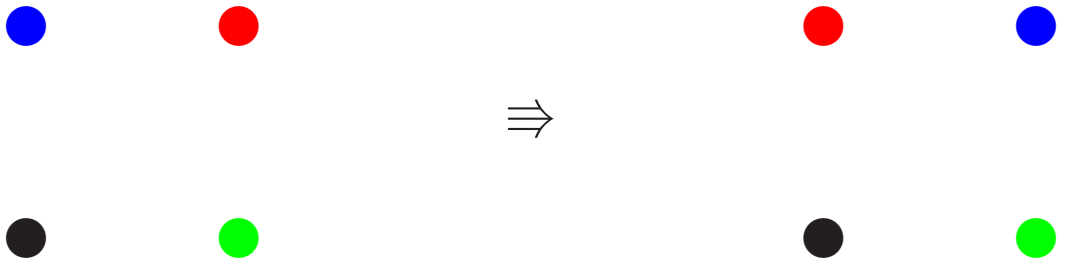
Let integer $1 \leqslant s \leqslant n$ be given. The system has a $(0,1)$-solution if and only if there is a $(0,1)$-solution to a new system in which for each $1 \leqslant j \leqslant m$ in the j-th equation the coefficient a_{js} of the variable x_s is replaced by a linear combination of the coefficients

$$c_j = 2b_j - \sum_{k=1}^{n} a_{jk}.$$

The geometric meaning is in the projective transformation, where: a hyperplane passing through the point $(2, \cdots, 2)$ in the affine space and not incident to any $(0,1)$-point maps to an improper hyperplane; $(0,1)$-points are mapped to, generally speaking, other $(0,1)$-points.

Such a transformation is an involution.

In the plane, $(0,1)$-points correspond to the vertices of a square.



Let us consider an auxiliary system in which a linear term from a new variable y is added to each equation:

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n + c_1 y &= b_1 \\ \cdots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n + c_m y &= b_m, \end{cases}$$

where the coefficients

$$c_j = 2b_j - \sum_{k=1}^{n} a_{jk}.$$

One of the solutions of this system is obtained when all variables are equal to 2.

The auxiliary system has a $(0,1)$-solution if and only if the original system has a $(0,1)$-solution. Moreover, the set of $(0,1)$-solutions to the auxiliary system is divided into pairs of antipodal solutions that transform into each other under simultaneous inversion of the values of all variables. Consequently, if a $(0,1)$-solution exists, then the auxiliary system has a pair of antipodal $(0,1)$-solutions for both $x_s = 0$ and $x_s = 1$. Next, let us fix the value $x_s = 0$ and replace the name of the variable y with x_s.

Lemma 2 is proven.

Theorem. *There is a polynomial-time algorithm that takes as input a system of m linear equations in n variables over the field $GF(3)$ and, subject to the condition*

$$m \leqslant \log_3 \log_3(2n-1),$$

accepts the input if and only if the system has a $(0,1)$-solution.

The input is an $m \times (n+1)$ matrix $M = [A \mid \mathbf{b}]$ of a system of equations $A\mathbf{x} = \mathbf{b}$. If the matrix A is empty, then the linear forms of all equations vanish, and the equations themselves turn into either identities $0 = 0$ or false equalities $0 = 1$ or $0 = 2$. The matrix M can be modified so that the new system of linear equations has a $(0,1)$-solution if and only if the original system of equations has a $(0,1)$-solution. In this case, the numbers of rows and columns never increase.
The loop executes the steps corresponding to Lemmata 1 and 2 until the matrix M stabilizes or the additional stopping condition is satisfied. For the resulting system of equations, the existence of a $(0,1)$-solution is easily verified if the conditions of the theorem are satisfied.

If there are k variables left, and the system contains several equations and cannot be reduced, then 2^k cases are analyzed. Let us estimate the number k from above.
Since the remaining number of equations does not exceed the number m, the inequality

$$\log_3(2k-1) < m$$

holds. But by the condition of applicability of the algorithm,

$$m \leqslant \log_3 \log_3(2n-1).$$

Therefore, the inequalities $(2k-1) < \log_3(2n-1)$ and

$$k \leqslant 0.5 \log_3(2n-1) < 0.3155 \log_2(2n-1)$$

hold. Therefore, the number of different $(0,1)$-values of the remaining k variables is less than the number $(2n-1)^{0.3155} = o(n)$.

Theorem is proven.

Example. The system of equations in two variables x_1 and x_2 over $GF(3)$

$$\begin{cases} 2x_1 + 2x_2 &= 1 \\ x_1 + 2x_2 &= 1 \end{cases}$$

corresponds to the augmented matrix

$$M = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 2 & 1 \end{pmatrix}.$$

There are no linearly dependent columns among the first two, but Lemma 2 can be applied. The column $\mathbf{c} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ is proportional to the first column in the matrix M. Replacing the second column, we obtain the matrix

$$\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix}.$$

Since the first and second columns are proportional to each other, it is possible to eliminate variables, which leads to the false equality $0 = 2$. Therefore, there are no $(0, 1)$-solutions.

For the number m of rows, the median and 99-th percentile are given for the number of randomly generated nonzero columns, none of which are linearly dependent, but the next column is linearly dependent on some of the previous columns. For each number m, $100,000$ series of columns are used. The upper bound on the number of pairwise independent columns is $(3^m - 1)/2$.

m	50%	99%	100%
4	7	18	40
8	67	172	3280
12	605	1561	265720
16	5450	14127	21523360
20	49176	126316	1743392200

The median of the largest number of columns in the series is close to the value

$$\frac{4}{5} \cdot \left(\frac{26}{15}\right)^m$$

The algorithm is implemented in Python.
The program either recognizes the presence of a $(0,1)$-solution or gives an indefinite answer.
To empirically estimate the running time, for different values of the number of equations m and the number of variables n, the medians of the running time **in seconds** were calculated under the condition of obtaining a definite answer, when the existence or absence of a $(0,1)$-solution was established.

	n			
m	10^5	10^6	10^7	10^8
4	0.03	0.3	3	35
8	0.09	0.9	10	97
12	0.17	1.8	18	187
16	0.33	2.9	28	304
20	0.9	4.7	43	445
24	2.06	10.4	63	609

The program as well as some examples are available at

`http://lab6.iitp.ru/-/havoc`

- The binary search method can find some $(0,1)$-solution to the system when one exists, although enumerating all $(0,1)$-solutions may be too difficult.

- This allows the possibility of practical use for solving those applied problems that can easily be reduced to finding a $(0,1)$-solution to a system of linear algebraic equations.

Thank you!