Probability of Unique Integer Solution to a System of Linear Equations

O. L. MANGASARIAN olvi@cs.wisc.edu
Computer Sciences Department
University of Wisconsin
Madison, WI 53706
Department of Mathematics
University of California at San Diego
La Jolla, CA 92093

BENJAMIN RECHT brecht@cs.wisc.edu
Computer Sciences Department
University of Wisconsin
Madison, WI 53706

Abstract. We consider a system of m linear equations in n variables $Ax = d$ and give necessary and sufficient conditions for the existence of a unique solution to the system that is integer: $x \in \{-1,1\}^n$. We achieve this by reformulating the problem as a linear program and deriving necessary and sufficient conditions for the integer solution to be the unique primal optimal solution. We show that as long as m is larger than $n/2$, then the linear programming reformulation succeeds for most instances, but if m is less than $n/2$, the reformulation fails on most instances. We also demonstrate that these predictions match the empirical performance of the linear programming formulation to very high accuracy.

Keywords: unique integer solution, linear equations, linear programming

1. INTRODUCTION

We consider the system of linear equations in the real vector variable x:

$$Ax = d, \tag{1}$$

where A is a given real matrix in $R^{m\times n}$, $d \in R^m$ and $x \in R^n$. We are interested in the conditions under which this system has a unique solution which is integer, that is

$$Ax = d, \ \ x \in \{-1,1\}^n. \tag{2}$$

This problem, which has also been studied in [12], can be considered a generalization of the classical knapsack feasibility problem [11, 7, 4] of finding an n-dimensional binary integer vector $y \in \{0,1\}^n$ such that:

$$a^T y = c, \tag{3}$$

where a is an n-dimensional column vector of positive integers and c is a positive integer. An obvious generalization of this is the following generalized multi-knapsack feasibility problem where there are no integrality or nonnegativity restrictions on the $m\times n$ real matrix A or the real vector $b \in R^m$:

$$Ay = b, \ \ y \in \{0,1\}^n. \tag{4}$$

Using the transformation:

$$y = \frac{e-x}{2},\ x = e - 2y, \tag{5}$$

where e is a column vector of ones, we obtain the absolute value equation [11, 16]:

$$\begin{array}{rcl} |x| & = & e, \\ Ax & = & d, \end{array} \tag{6}$$

where:

$$d = Ae - 2b. \tag{7}$$

It is evident then that (6) is equivalent to our original problem (2).

Unfortunately, even if an integer solution is provided, determining the uniqueness of a given integer solution of a problem such as (2) is an NP-hard problem [17, 15, 14]. To circumvent this difficulty, we provide necessary and sufficient conditions that (1) has a unique solution in the hypercube $[-1,1]^n$ which in turn is integer. We shall do this in Section 2 by solving a linear programming problem. In Section 3 we give the probability that a randomly generated solvable problem (2) will indeed have a unique integer solution. In particular we show that as long as the number of rows is greater than half the number of columns, then for most equations of the form (2) which have an integer solution, the corresponding integer solution is unique and can be computed via linear programming. A related probabilistic result is obtained in [3] that utilizes a face counting technique in contrast to our simple linear programming uniqueness approach here. In Section 4 we shall give some numerical examples illustrating our results and shall conclude the paper in Section 5.

A word about our terminology and notation now. When we refer to an integer solution x of either the linear equation (1) or the linear program (10) below, we mean exactly that $x \in \{-1,1\}^n$ and exclude the case when a component of x is zero. All vectors will be column vectors unless transposed to a row vector by a superscript T. For a vector $x \in R^n$ the notation x_j will signify the j-th component, $|x|$ denotes the vector in R^n whose components are the absolute values of the components of x, and $\|x\|_p$ denotes the p-th norm of x. The scalar (inner) product of two vectors x and y in the n-dimensional real space R^n will be denoted by $x^T y$. The notation $A \in R^{m\times n}$ will signify a real $m \times n$ matrix. For such a matrix, A^T will denote the transpose of A, A_i will denote the i-th row and A_{ij} the ijth element. A vector of ones in a real space of arbitrary dimension will be denoted by e. Thus for $e \in R^n$ and $x \in R^n$ the notation $e^T x$ will denote the sum of the components of x. A vector of zeros in a real space of arbitrary dimension will be denoted by 0. The abbreviation "s.t." stands for "subject to".

2. Linear Programming Formulation and Uniqueness of Solution that is Integer

Our analysis is based on the observation that if x_0 is integer and is the unique solution of $Ax = d$ in the hypercube $[-1,1]^n$, then x_0 is the unique solution in $\{-1,1\}^n$ as well. Finding a solution in $[-1,1]^n$ can be reduced to linear programming, and, moreover, we can readily provide necessary and sufficient conditions that the resulting solution is unique.

Since the hypercube $[-1,1]^n$ is equal to the unit ball $\|x\|_\infty \le 1$ of the ℓ_∞ norm:

$$\|x\|_\infty = \max_{1\le i\le n} |x_i|, \tag{8}$$

we can try to find a solution in $[-1,1]^n$, by solving the ℓ_∞ norm minimization problem

$$\min_x \|x\|_\infty \;\; s.t. \;\; Ax = d. \tag{9}$$

Suppose that there exists an $x_0 \in \{-1,1\}^n$ satisfying $Ax = d$. Under what conditions is x_0 also the unique optimal solution of (9)?

Note that we can reformulate (9) as the following linear program

$$\min_{x,\delta} \delta \;\; s.t. \;\; Ax = d, \;\; -\delta e \le x \le \delta e. \tag{10}$$

The dual problem of (10) is given by

$$\max_{u,v,w} d^T u \;\; s.t. \;\; A^T u - v + w = 0, \;\; e^T(v+w) = 1, \;\; (v,w) \ge 0. \tag{11}$$

Note also that the primal linear program (10) is a convex relaxation of the absolute value equation (6) where we replace the first equality of (6) by two inequalities, and we replace the right hand side e with a variable δe which we attempt to minimize. With this reformulation, we can use the necessary and sufficient conditions of [9] to verify that x_0 is a unique solution of the linear program (10):

THEOREM 1 **[9, Theorem 2(iii)]** *Let $\bar{x}$ be a solution of the linear program*

$$\min_x h^T x; \;\; s.t. \;\; Gx = g, \;\; Px \ge q. \tag{12}$$

Let P_{eq} denote the submatrix of P consisting of the rows of $Px \ge q$ for which $P_i\bar{x} = q_i$. Then $\bar{x}$ is unique if and only if there exists no z satisfying

$$Gz = 0, \;\; P_{\mathrm{eq}} z \ge 0, \;\; h^T z \le 0, \;\; z \ne 0 \tag{13}$$

With this in hand, we can state now our principal result.

Proposition 2 Uniqueness of Solution of (1) that is Integer *A necessary and sufficient condition that the linear program (10) has a unique integer solution is that it has a minimum value of* 1 *with solution $x \in \{-1,1\}^n$ such that for the diagonal matrix D of ± 1's defined as:*

$$D = \mathrm{diag}(x) \tag{14}$$

the system:

$$DA^T r > 0, \tag{15}$$

has a solution $r \in R^n$.

Proof: The constraints of the linear program (10) imply that $\|x\|_\infty \le \delta$. Hence a necessary and sufficient condition for $x \in \{-1,1\}^n$ to be a solution of (10) is that the corresponding minimum value of the objective function δ is 1. It follows from Theorem 1 above that $x \in \{-1,1\}^n$ is a unique solution of (10), if and only if the following holds for the diagonal matrix of ± 1's $D = \mathrm{diag}(x)$:

$$As = 0,\ \ Ds + e\delta \geq 0,\ \ \delta \leq 0,\ \ \text{has no solution}\ \ (s,\delta) \neq 0. \tag{16}$$

Define now $z = Ds$, and since $DD = I$, we also have that $s = Dz$. Hence condition (16) is equivalent to:

$$ADz = 0,\ \ z + e\delta \geq 0,\ \ \delta \leq 0,\ \ \text{has no solution}\ \ (z,\delta) \neq 0. \tag{17}$$

We can eliminate δ by reformulating (17) as follows:

$$ADz = 0,\ \ z \geq 0,\ \ \text{has no solution}\ \ z \neq 0. \tag{18}$$

To see that (18) is equivalent to (17), observe that if $ADz = 0,\ z + e\delta \geq 0,\ \delta \leq 0$ and $(z,\delta) \neq 0$, then $z \geq -e\delta \geq 0$ implying that $z \neq 0$, for otherwise δ would also equal zero. Conversely, if $ADz = 0$, $0 \neq z \geq 0$, then for $\delta = 0$ we have that $(z,\delta) \neq 0$ and $z + e\delta \geq 0$.

Now, by using Gordan's theorem of the alternative [10, Theorem 2.4.5], condition (18) is equivalent to $DA^T r > 0$ having a solution r, which is the desired condition (15) ∎

Note that Proposition 2 can be easily implemented by solving the linear program (10) and checking that its minimum objective function value is $\delta = 1$. Also, if there exists a dual optimal solution with $d'u = \delta = 1$ which has the property that for all i either v_i or w_i is strictly positive, then it follows from the complementarity conditions: $v^T(-x+e) = 0$ and $w^T(x+e) = 0$ for the primal optimal solution x, that $x \in \{-1,1\}^n$. The search for such a dual optimal solution can also be accomplished by linear programming.

We also note here that a somewhat different linear programming uniqueness characterization [1] can be employed to obtain different uniqueness conditions than those of (14)-(15) above. However our condition (15), which is equivalent to that of the columns of the matrix AD lying in the same hemisphere of R^m, is key in deriving our probalistic results of Section 3.

We now proceed to give conditions that the linear program (10) returns a unique integer solution for problem (2) with a likely probability.

3. Probability that the Linear Program (10) Solves the Integer Problem (2)

While the conditions in Proposition 2 are completely deterministic and checkable, we have not yet shown that there exist matrices A satisfying these conditions. In this section, we show that as long as the ratio m/n is greater than $1/2$, then we can solve the integer programming problem for "most" A by solving the linear program (10).

The existence of an $r \in R^m$ satisfying $DA^T r > 0$ is simply equivalent to the columns of the matrix AD lying in the same hemisphere of R^m. We now quantify a very general family of random matrices for which we can precisely calculate the probability that such an event occurs. We say that A is a *generic random matrix* if all sets of m columns of A are linearly independent with probability 1 and that each column of A is symmetrically distributed about the origin. Wendel [18] showed via a simple inductive argument that the probability of all of the columns of a generic random matrix lying in the same hemisphere is precisely equal to

$$p_{m,n} = 2^{-n+1} \sum_{i=0}^{m-1} \binom{n-1}{i}. \tag{19}$$

If A is a generic random matrix, then so is AD, and it follows that the probability that (10) has a unique integer solution which is recovered by the ℓ_∞ norm heuristic is *exactly* $p_{m,n}$.

It is rather surprising that not only can we can compute the probability of uniqueness in closed form for this problem, but that it is equal to the probability that at most m-1 heads appear in n-1 fair coin tosses. It is easy to check that for a fixed n, $p_{m,n}$ is an increasing function of m and that:

$$p_{1,n} = 2^{-n+1}, \;\; p_{m,2m} = \frac{1}{2}, \;\; p_{n,n} = 1, \tag{20}$$

where the last two equalities are easily obtained by elementary properties of binomial coefficients. Moreover, we can use standard tail bounds of the binomial distribution to describe asymptotically when (10) has a unique solution. For instance, if we set $\gamma = (m-1)/(n-1)$, then Hoeffding's inequality [6] states that

$$\begin{aligned} p_{\gamma n,n} &\geq 1 - \exp(-2(\gamma - 1/2)^2(n-1)) \qquad \gamma > 1/2 \\ p_{\gamma n,n} &\leq \exp(-2(\gamma - 1/2)^2(n-1)) \qquad \gamma < 1/2\,. \end{aligned} \tag{21}$$

That is, for a fixed ratio γ, the probability that the heuristic yields a unique integer solution goes to 1 exponentially with n for $\gamma > 1/2$, and the probability that the heuristic yields a unique integer solution goes to 0 exponentially with n for $\gamma < 1/2$. Our computational results of the next section will support these facts.

As a final note, we can use Wendel's theorem to count the number of integer x_0 that can be recovered via ℓ_∞ minimization. Suppose that all subsets of m columns of the matrix A are linearly independent. For how many $x_0 \in \{-1,1\}^n$ does it hold that x_0 is the unique integer solution of $Ax = Ax_0$? The answer is exactly $2^n p_{m,n}$. We can prove this probabilistically by letting x_0 be sampled uniformly from $\{-1,1\}^n$. By Proposition 2, x_0 is the unique integer solution of $Ax = Ax_0$ if and only if there exists an $r \in R^m$ with $\operatorname{diag}(x_0)A^T r > 0$. Since $A \operatorname{diag}(x_0)$ is a generic random matrix, the probability of the existence of such an r is $p_{m,n}$ which proves our assertion. This means that for most generic random matrices A, our ℓ_∞ norm heuristic will succeed for most of the possible d vectors in (2) as long as $m/n > 1/2$.

4. Computational Results

We tested our linear programming formulation (10) by running it on randomly generated linear integer equations (2). We summarize our computational results as follows.

In Table 1 we present average computational results for 10 runs for each of 9 cases of solvable integer linear equations (2) solved by the linear program (10), utilizing the CPLEX linear programming code [8] within MATLAB [13]. We generated the $m \times n$ matrix A containing pseudorandom values drawn from the standard normal distribution. The right

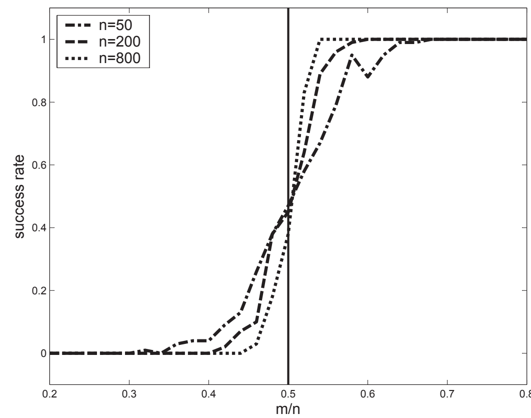

Figure 1. Probabilities that the uniqueness of the integer solution of the system of linear equations (2) $Ax = d$ by the linear program (10), utilizing the CPLEX linear programming code [8] within MATLAB [13]. For each n, we selected various values of m and ran 100 experiments. We declare a success if the returned solution equalled the true integer solution. Empirical rates are plotted here for $n = 50$, 200, and 800. The solid vertical line denotes the predicted phase transition where $m/n = 1/2$.

hand side d of (2) was set equal to Az where each component of z was set to 1 or -1 with equal probability. The average times in column 3 of Table 1 are for a 4 Gigabyte machine running Red Hat Enterprise Linux 5. Column 4 gives the average minimum over 10 cases of the minimum value of δ of the linear program (10) which indicates a unique integer solution of (2) when it is equal to 1. Column 5 of Table 1 gives the number of runs out of 10 that the linear program (10) returned an integer solution of (1). We make the following remarks regarding Table 1.

(i) We note that for all cases for which $p_{m,n} > 1/2$, that is cases for which $m > n/2$, the linear program (10) returned an exact integer solution.

(ii) Out of the 30 cases for which $m = n/2$, exactly 14 linear programs (10) returned integer solutions of (2). This is in remarkable agreement with the probability of $p_{m,2m} = \frac{1}{2}$ given above in (20).

For a graphic display of solution behavior, we ran numerous experiments for the cases $n = 50$, 200, and 800. For each pair of m and n, we tested 100 random instances and declared success if the optimal solution was equal to the generated z as defined in the previous paragraph. As depicted in Figure 1, there is a dramatic transition between failure and success of the heuristic as the ratio m/n increases. This transition is exactly predicted by the results of Section 3. The solid vertical line in the plot is the predicted phase transition where the probability of success is computed to be $1/2$. As n grows, the shape of this curve rapidly approaches a step function equal to 0 for $m/n < 1/2$ and 1 for $m/n > 1/2$.

Table 1. Integer solution of the system of linear equations (2) $Ax = d$ by the linear program (10), utilizing the CPLEX linear programming code [8] within MATLAB [13]. Each line in the first four columns represents the average of ten runs. The times are for a 4 Gigabyte machine running Red Hat Enterprise Linux 5. Column 4 gives the average minimum value of the objective function δ of the linear program (10), which indicates a unique integer solution of (2) when it is equal to 1.

No. of Rows m	No. of Variables n	MATLAB Time Sec toc	Minimum Value of δ	No. of Runs Out of 10 Returning an Integer Solution
250	500	0.4420	0.9953	6
300	500	0.4290	1	10
400	500	0.4340	1	10
500	1,000	6.1820	0.9950	4
600	1,000	5.4240	1	10
800	1,000	4.2810	1	10
750	1,500	37.3870	0.9957	4
900	1,500	47.3180	1	10
1,200	1,500	19.3550	1	10

5. Conclusion and Outlook

We have presented a method to transform an integer programming problem into a linear program, which under appropriate conditions, yields a unique integer solution to the integer program. Using this formulation we have been able to analyze random instances of the integer program and classify which instances are readily solvable in polynomial time with high probability.

In some sense, a popular body of work in compressed sensing follows a similar trajectory (see, for instance [5, 2]). There, an NP-Hard problem of finding the sparsest solution to $Ax = b$ is replaced by a linear program, and a dual certificate is produced to guarantee uniqueness. The existence of such a certificate is then guaranteed by appealing to probabilistic arguments. In the compressed sensing literature, this certificate is sufficient, but not necessary for the linear programming solution to coincide with the sparsest solution. It would be interesting to extend our linear programming techniques to provide necessary and sufficient conditions for optimality in compressive sensing and other NP-HARD optimization problems.

Acknowledgements

Research described here is available as Data Mining Institute Report 09-02, September 2009: ftp://ftp.cs.wisc.edu/pub/dmi/tech-reports/09-02.pdf. The authors wish to thank their colleagues Eric Bach and Ali Rahimi for helpful discussions about the topic of this paper.

References

1. G. Appa. On the uniqueness of solutions to linear programs. *Journal of the Operational Research Society*, 53:1127–1132, 2002.
2. E. J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, 2006.
3. David L. Donoho and Jared Tanner. Counting the faces of randomly-projected hypercubes and orthants with applications. *Discrete and Computational Geometry*, 43(3):522–541, 2010.
4. D. Fayard and G. Plateau. An algorithm for the solution of the 0–1 knapsack problem. *Computing*, 28:269–287, 1982.
5. J. J. Fuchs. On sparse representations in arbitrary redundant bases. *IEEE Transactions on Information Theory*, 50:1341–1344, 2004.
6. W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.
7. R. Horst, P. Pardalos, and N. V. Thoai. *Introduction to Global Optimization*. Kluwer Academic Publishers, Dodrecht, Netherlands, 1995.
8. ILOG, Incline Village, Nevada. *ILOG CPLEX 9.0 User's Manual*, 2003. http://www.ilog.com/products/cplex/.
9. O. L. Mangasarian. Uniqueness of solution in linear programming. *Linear Algebra and Its Applications*, 25:151–162, 1979.
10. O. L. Mangasarian. *Nonlinear Programming*. SIAM, Philadelphia, PA, 1994.
11. O. L. Mangasarian. Knapsack feasibility as an absolute value equation solvable by successive linear programming. *Optimization Letters*, 3:161–170, 2009. ftp://ftp.cs.wisc.edu/pub/dmi/tech-reports/08-03.pdf. Optimization Letters online version: http://www.springerlink.com/content/7011287432285747/.
12. O. L. Mangasarian and M. C. Ferris. Uniqueness of integer solution of linear equations. Technical Report 09-01, Data Mining Institute, Computer Sciences Department, University of Wisconsin, Madison, Wisconsin, July 2009. ftp://ftp.cs.wisc.edu/pub/dmi/tech-reports/09-01.pdf. Optimization Letters 4, 2010, 559-565.
13. MATLAB. *User's Guide*. The MathWorks, Inc., Natick, MA 01760, 1994-2006. http://www.mathworks.com.
14. C.H. Papadimitriou. On the complexity of unique solution. In *Proceedings of the 23rd Annual Symposium on Foundations of Computer Science*, pages 14–20, Chicago, IL, 1982. IEEE CNF.
15. O. A. Prokopyev, H.-X. Huang, and P. M. Pardalos. On complexity of unconstrained hyperbolic 0-1 programming problems. *Operations Research Letters*, 33:312–318, 2005.
16. J. Rohn. A theorem of the alternatives for the equation $Ax + B|x| = b$. *Linear and Multilinear Algebra*, 52(6):421–426, 2004. http://www.cs.cas.cz/ rohn/publist/alternatives.pdf.
17. L. G. Valiant and V. V. Vazirani. Np is as easy as detecting unique solutions. In *Annual ACM Symposium on Theory of Computing, Proceedings of the 17th Annual ACM Symposium on Theory of Computing*, pages 458–463, New York, NY, 1985. Association for Computing Machinery.
18. J. G. Wendel. A problem in geometric probability. *Mathematica Scandinavica*, 11:109–111, 1962.