

NEW RESULTS ON THE AVERAGE BEHAVIOR OF SIMPLEX ALGORITHMS¹

BY ILAN ADLER, NIMROD MEGIDDO² AND MICHAEL J. TODD³

ABSTRACT. It has been a challenge for mathematicians to theoretically confirm the extremely good performance of simplex algorithms for linear programming. We have confirmed that a certain variant of the simplex method solves problems of order $m \times n$ in an expected number of steps which is bounded between two quadratic functions of the *smaller* dimension of the problem. Our probabilistic assumptions are rather weak.

1. Introduction. We consider the linear programming problem of order $m \times n$ in the form: Maximize $c^T x$ over all $x \geq 0$ in R^n such that $Ax \leq b$, where $A \in R^{m \times n}$. It has been observed that the simplex algorithms for linear programming, developed by George Dantzig [D], work extremely well. A central question in the field of analysis of algorithms is to estimate the expected number of steps that these algorithms perform relative to different probability distributions of inputs. Some background is given in §2. In [AM1 and T] we consider a variant of the simplex method called the “lexicographic self-dual method”. We find that the expected number of steps of this variant, relative to a rather weak probabilistic model, is only $O((\min(m, n))^2)$. This is the first polynomial average-case upper bound for a simplex algorithm which is capable of solving any linear programming problem. The first two authors [AM2] have also determined a quadratic lower bound, so the behavior of this variant is indeed quadratic, whereas it has been conjectured that other variants take a linear expected number of steps. With various choices of the starting point, this variant can simulate different “constraint-by-constraint” and “variable dimension” algorithms, as well as combinations thereof [Me2]. Adler, Karp and Shamir [AKS] provide a different proof of similar upper bound for a certain class of “constraint-by-constraint” algorithms. It follows from [Me2] that their variant can also be simulated by the self-dual algorithm with an appropriate starting point.

2. Background. The simplex algorithms iteratively change the basis of a linear system of equations, until they reach an “optimal” basis, or a basis that exhibits that no optimal solution exists. Potentially, there are $\binom{m+n}{m}$

Received by the editors February 3, 1984 and, in revised form, June 20, 1984.

1980 *Mathematics Subject Classification.* Primary 68C25; Secondary 90C05.

¹This is an announcement of the results in [AM1, AM2, T]. Copies are available from the authors.

²The work of this author was supported in part by the National Science Foundation under Grants MCS-8300984, ECS-8218181 and ECS-8121741.

³The work of this author was supported in part by a fellowship from the Alfred P. Sloan Foundation and by NSF Grant ECS-8215361.

©1984 American Mathematical Society
0273-0979/84 \$1.00 + \$.25 per page

bases, so this is an obvious upper-bound on the number of steps that any simplex algorithm can perform. Klee and Minty [KM] and, later, others have shown that many simplex algorithms require an exponential number of steps in the worst case. A bad example for the variant discussed here was given by Murty [Mu]. Computational experience has shown that the number of steps is usually much smaller. This has been observed in practice and in controlled experiments referenced in our detailed reports. It has been a challenge to confirm these findings theoretically. Note that the ellipsoid algorithm for linear programming [K], which has a worst-case bound that is polynomial in m , n and the logarithms of the numbers involved, appears at present to be inefficient for practical solution of linear programming problems.

Borgwardt [Bo1, Bo2] and Smale [S1, S2] have recently provided probabilistic analyses of simplex algorithms. Both assume that the vectors generating the problem are sampled from spherically symmetric distributions; however, Smale actually obtains his results under a weaker model of symmetry. Borgwardt analyzes a *subclass* of the general class of linear programming problems. He shows that the expected number of steps is $O(n^4 m^{1/(n-1)})$. We note that this upper bound tends to infinity when either m or n tend to infinity. Smale considers the so-called self-dual simplex algorithm [D]. He shows that the expected number of steps is less than $c(m)(\ln n)^{m(m+1)}$. Obviously, this upper bound tends to infinity with n . Blair [Bl] proves an upper bound of $c(m)(\ln n)^{m(m+1)\ln(m+1)+m}$ under a weaker model. Megiddo [Mel] has shown that under Smale's model the expected number of steps decreases to a limit $c(m)$ whenever n tends to infinity while m is fixed.

Our papers [AM1, AM2, T] considerably improve upon the previous results. We confirm that the average number of steps is polynomial in the *smaller* dimension of the problem. We analyze the average number of pivot steps performed by a variant of the self-dual simplex algorithm described below. This variant yields a much better estimate of the number of steps. We show that this number is bounded between two quadratic functions of the *minimum* of the two dimensions. Furthermore, our analyses in these papers are carried out under models which are much weaker than Borgwardt's and Smale's. Instead of complete spherical symmetry, we require only symmetry with respect to certain reflections, together with a certain regularity condition on the matrix with probability one, which holds if the problem is sampled from any continuous distribution. We note that Haimovich [H] and Adler [A] consider the weak probability model we describe below and independently obtain linear bounds on the expected length of an efficient path determined by two objective functions; however, these results do not explain the behavior of an algorithm for solving a general linear programming problem.

3. The probabilistic model. Our results are valid for *any* distribution in a wide class which we define below. Natural probabilistic models to look at are those with some symmetry assumptions. It is desirable to have the group of symmetries as small as possible. Under the spherically symmetric model each class contains a continuum of instances.

Under our model the classes are finite. Given an instance (A, b, c) , let $A^* \in R^{(m+1) \times (n+1)}$ denote a matrix consisting of A together with b as an additional column and c^T as an additional row. There is thus a redundant entry $A_{m+1, n+1}^*$. It is convenient to assume every submatrix of A^* is nonsingular with probability one, which is true if A^* is sampled from any continuous distribution.

We first describe the weaker model under which we prove the upper bound result. The distribution is required to be invariant under the group of symmetries generated by the $m + n$ transformations of multiplying either one of the first n columns, or one of the first m rows of the matrix A^* by -1 . Thus, in every equivalence class, each of the 2^{m+n} members is equally probable. It is interesting to mention that the number of symmetries cannot be subexponential if we are to prove a polynomial upper bound on the average number of steps, since in the worse case the number is exponential.

The stronger model, under which we are able to prove the lower bound result, requires that all entries of A^* be independent, identically distributed random variates whose common distribution is symmetric about the origin.

4. The algorithm. It is known from the theory of linear programming that $x \in R^n$ (such that $Ax \leq b$ and $x \geq 0$) is an optimal solution if and only if there exists $y \in R^m$ such that $y^T A \geq c^T$, $y \geq 0$ and the following equalities (called “complementary slackness” conditions) hold: $y^T(Ax - b) = (y^T A - c^T)x = 0$. The problem may be restated in the following way: Given (A, b, c) , find $(x, y, u, v) \geq 0$ such that $Ax + u = b$, $v - A^T y = -c$ and $y^T u = v^T x = 0$. Since the variables are “complementary” (in the sense that $x_j v_j = 0$, $j = 1, \dots, n$, and $y_i u_i = 0$, $i = 1, \dots, m$), we can use a vector $\xi \in R^{m+n}$ to describe all of them as follows. We identify $\xi_j^+ = v_j$ and $\xi_j^- = -x_j$ for $j = 1, \dots, n$, and $\xi_{n+i}^+ = u_i$ and $\xi_{n+i}^- = -y_i$ for $i = 1, \dots, m$. Equivalently, the problem is: Find ξ such that $F(\xi) = q$, where F is a certain piecewise linear mapping from R^{m+n} into itself and $q = (-c^T, b^T)^T$. The mapping F is linear on each orthant and coincides with the identity on the positive orthant. Obviously, if q is positive then $F(q) = q$ and the problem is solved in a trivial way. In general, the algorithm can start at any *positive* point q_0 and follow $F^{-1}((1-t)q_0 + tq)$, while t goes from 0 to 1. When the inverse image crosses the boundary between two orthants of R^{m+n} , a “pivot” step occurs. Thus, the essence of the analysis is to estimate the number of times the inverse image crosses such boundaries. The self-dual method uses the starting point $q_0 = (1, \dots, 1)^T$. It turns out that the analysis is much more favorable and tractable when q_0 consists of pairwise distinct powers of ε , where ε is sufficiently small. Also, in that case the sequence of orthants visited by the inverse image of the segment $[q_0, q]$ can be generated by lexicographic pivoting rules.

5. Outline of the analysis. The key to estimating the expected number of steps is the evaluation of the probabilities of bases occurring in the solution process. These events were characterized by Smale in terms of inclusions of random vectors in random cones in R^{m+n} . The random vector is q . The cones are spanned by columns and rows of A (augmented with zeros), a few

unit vectors and the vector q_0 . Alternatively, the roles of q and q_0 can be reversed. There are $\binom{m+n}{m}$ cones to consider. For each cone, the bounds obtained depend on two indices, namely, the smallest powers of ε which occur in two critical sections of the vector q_0 . Once these bounds are obtained (all are powers of $1/2$), it only remains to count the number of cones with the same indices and sum up the probabilities. The reader is referred to the detailed reports for the complete analysis.

To illustrate how the probabilities are estimated, we give an analogous simplistic example. Let $u \in R^k$ be fixed and let $S \in R^{k \times k}$ be generated as is A^* in the (weaker) model of §3. Then the probability that u lies in the cone $C(S)$ spanned by the columns of S is 2^{-k} . This easily follows by considering all 2^k matrices S' obtained from S by reversing the signs of some of its columns; with probability one, u lies in $C(S')$ for just one such S' . Our analysis is complicated by the unit vectors, the form of q_0 , and the extra zeroes.

REFERENCES

- [A] I. Adler, *The expected number of pivots needed to solve parametric linear programs and the efficiency of the self-dual simplex method*, Dept. Industrial Engineering and Operations Res., Univ. of California, Berkeley, June 1983.
- [AKS] I. Adler, R. M. Karp and R. Shamir, *A simplex variant solving an $m \times d$ linear program in $O(\min(m^2, d^2))$ expected number of steps*, Report UCB CSD 83/158, Computer Science Division, Univ. of California, Berkeley, December 1983.
- [AM1] I. Adler and N. Megiddo, *A simplex-type algorithm solves linear programming problems of order $m \times n$ in only $O((\min(m, n))^2)$ steps on the average*, November 1983 (manuscript).
- [AM2] ———, *A simplex algorithm whose average number of steps is bounded between two quadratic functions of the smaller dimension*, Proc. 16th Annual ACM Sympos. on Theory of Computing, 1984, pp. 312–323.
- [Bl] C. E. Blair, *Random linear programs with many variables and few constraints*, Faculty Working Paper No. 946, College of Commerce and Business Administration, Univ. of Illinois at Urbana-Champaign, April 1983.
- [Bo1] K.-H. Borgwardt, *Some distribution-independent results about the asymptotic order of the average number of pivot steps of the simplex method*, Math. Oper. Res. **7** (1982), 441–462.
- [Bo2] ———, *The average number of steps required by the simplex method is polynomial*, Z. Oper. Res. Ser. A–B **26** (1982), 157–177.
- [D] G. B. Dantzig, *Linear programming and extensions*, Princeton Univ. Press, Princeton, N.J., 1963.
- [H] M. Haimovich, *The simplex algorithm is very good!—On the expected number of pivot steps and related properties of random linear programs*, Columbia Univ., New York, April 1983 (manuscript).
- [KM] V. Klee and G. J. Minty, *How good is the simplex algorithm?*, Inequalities. III, Academic Press, 1972, pp. 159–175.
- [K] L. G. Khachian, *A polynomial algorithm in linear programming*, Soviet Math. Dokl. **20** (1979), 191–194.
- [Me1] N. Megiddo, *Improved asymptotic analysis of the average number of steps performed by the self-dual simplex algorithm*, Dept. of Computer Science, Stanford Univ., September 1983 (preliminary report).
- [Me2] ———, *A note on the generality of the self-dual algorithm with various starting points*, Dept. of Computer Science, Stanford Univ., December 1983 (preliminary report).
- [Mu] K. G. Murty, *Computational complexity of parametric linear programming*, Math. Programming **19** (1980), 213–219.

[S1] S. Smale, *On the average number of steps of the simplex method of linear programming*, Math. Programming **27** (1983), 241–267.

[S2] ———, *The problem of the average speed of the simplex method*, Mathematical Programming: The State of the Art (A. Bachem, M. Grötschel and B. Korte, eds.), Springer-Verlag, 1983, pp. 530–539.

[T] M. J. Todd, *Polynomial expected behavior of a pivoting algorithm for linear complementarity and linear programming problems*, Tech. Rep. No. 595, School of Operations Res. and Industrial Engineering, Cornell Univ., 1983.

DEPARTMENT OF INDUSTRIAL ENGINEERING AND OPERATIONS RESEARCH, UNIVERSITY OF CALIFORNIA, BERKELEY, CALIFORNIA 94720 (Current address of I. Adler)

SCHOOL OF OPERATIONS RESEARCH AND INDUSTRIAL ENGINEERING, CORNELL UNIVERSITY, ITHACA, NEW YORK 14853 (Current address of M. Todd)

Current address (N. Megiddo): IBM Research, 5600 Cottle Road, San Jose, California 95193

DEPARTMENT OF COMPUTER SCIENCE, STANFORD UNIVERSITY, STANFORD, CALIFORNIA 94305

XEROX PALO ALTO RESEARCH CENTER, PALO ALTO, CALIFORNIA 94304