

The basic George B. Dantzig, by Richard W. Cottle, Stanford University Press, Stanford, California, 2003, xvi + 378 pp., hardcover, \$57.00, ISBN 978-0-8047-4834-6

The modern study of optimization began with the introduction of the famous simplex method for linear programming by George Dantzig in 1947. Its growth since then has been remarkable: extensions to integer and combinatorial optimization, to quadratic programming and linear complementarity problems, to stochastic programming, and to conic, convex, and more general nonlinear optimization.

Even more surprising is that through the last sixty years, the simplex method has remained a powerful algorithm for the solution of linear programming problems of a scale at least four orders of magnitude larger than it was originally developed for. Indeed, problems with millions of variables, given suitable sparsity and structure, can be solved almost routinely on personal computers. For the last twenty-five years, the simplex method has had to share the spotlight with interior-point methods inspired by Karmarkar's projective algorithm, announced with considerable fanfare in 1984. And yet the simplex method has remained, if not the method of choice, a method of choice, usually competitive with, and on some classes of problems superior to, the more modern approaches.

In view of its unusual longevity, it is worth mentioning that the method was almost stillborn. The feasible region of a linear optimization problem is a convex polyhedron, the solution set to a system of linear inequalities. If the polyhedron is pointed (contains no lines), any linear objective function that attains its optimal value over such a set attains it at a vertex. The simplex method proceeds from vertex to vertex, always improving the objective function, until an optimum is reached (or unboundedness is demonstrated). Viewed this way, the method seems likely to be very inefficient: consider a feasible region that resembles a multi-faceted disco ball in three dimensions. Fortunately, Dantzig also considered another geometric point of view, where a problem is viewed in the space of possible right-hand sides rather than the space of the variables. From this perspective, the method appeared much more promising, enough so that it was tested and found effective for the small models that could be treated at that time. As Dantzig later said, "Our intuition in higher dimensions isn't worth a damn."

Prior to 1947, work related to linear programming was remarkably sparse. Much attention had been paid to systems of linear equations, with computational methods of Gauss and Jordan. Similarly, the study of certain optimization problems, sometimes with equality constraints, had been considered, for example, by Gauss, Euler, and Lagrange. Work on linear inequalities was much rarer: Fourier and de la Vallée Poussin had considered methods (basically the simplex method) to find solutions to a particular system of linear inequalities, and various alternative theorems had been established by Gordan, Farkas, and others. But general linear optimization subject to inequality constraints had been ignored, partly because of its difficulty and partly because of a lack of perceived (physical) applications.

2010 *Mathematics Subject Classification*. Primary 01A60, 65K05, 90-03, 90Cxx.

The need for systematic methods to analyze and improve complex systems after World War II changed these perceptions. Dantzig was working for the U. S. Air Force and was asked to develop models and methods for large-scale planning (or “programming” in military jargon) to coordinate the use of manpower and resources to fulfill military goals. His influences included work by W. Leontief in building large activity analysis or input-output models of the U.S. economy. Also relevant was a model by J. von Neumann on optimal growth of an economy. Finally, Dantzig’s own Ph.D. dissertation in statistics under J. Neyman had used ideas related to linear semi-infinite optimization and had employed a geometry that enabled him later to see the simplex method as a potentially efficient algorithm. Using these ideas, Dantzig was able to formulate the feasibility constraints for his plans using linear equations and inequalities. He also appreciated the need to choose a particular plan from all those feasible by minimizing or maximizing a linear objective. Such a problem became known as a linear programming problem, following a suggestion of T. C. Koopmans, who was hugely influential in spreading the new ideas among a generation of economists. The late 1940s was an ideal time for these ideas to germinate rapidly. The needs of the military to solve large logistics problems, the interest of economists in understanding large models of the economy, and the advent or promise of the electronic computer all combined to accelerate the development of optimization theory and algorithms over the following decades, assisted by burgeoning commercial applications in oil and other industries. One should not ignore developments in the Soviet Union, which might have been thought ideally suited to such tools for central economic planning. L. V. Kantorovitch developed ideas very similar to Dantzig’s in 1939, but his proposals fell on deaf ears: Soviet economists were suspicious of noncentrally planned prices and of mathematical intrusions into their domain in general.

During the 1950s and 1960s, optimization developed at a very rapid pace. A far-from-complete listing follows: von Neumann and A. W. Tucker and his students D. Gale and H. W. Kuhn developed the theoretical underpinnings of duality; extensions to combinatorial and integer optimization were made by M. M. Flood, L. R. Ford and D. R. Fulkerson, A. J. Hoffman, H. W. Kuhn, R. E. Gomory, and W. L. Eastman; to nonlinear programming by Kuhn and Tucker among others; to stochastic programming by E. M. L. Beale and Dantzig; to large-scale programming by Dantzig and P. Wolfe; and to complementary pivot theory by R. W. Cottle, Dantzig, and C. E. Lemke. The term “mathematical programming” was devised to include all these optimization problems. As is somewhat apparent from this very brief list, Dantzig was highly influential in a great number of these extensions, and remained very productive through the 1980s and beyond.

As well as being a discipline of great applicability in engineering, science, industry, and government, optimization has inspired much fascinating research in mathematics and has been enriched by many surprising areas of mathematics. Let us start with the simplex method for linear programming, which, as mentioned above, proceeds from vertex to vertex of the feasible region. While the scale of linear programming problems has grown from those with tens of constraints to those with hundreds of thousands of constraints, one remarkable empirical fact has remained true: in almost all cases, the number of steps to solve a problem with m equality constraints in n nonnegative variables (this is a universal form of the problem) is almost always at most a small multiple of m , say $3m$. This behavior of the method is responsible for its great efficiency in practice, and raises many questions:

What is the combinatorial structure of convex polyhedra? What bounds can be proved on their diameters? Are there linear programming problems requiring exponentially many steps? And since there are many variants of the simplex method distinguished by their rules for selecting the next vertex (so-called pivot rules), is there a pivot rule requiring only polynomially many steps?

A classic book on convex polyhedra is that of B. Grünbaum [12]; a more recent monograph, attuned more to the questions above, is that of G. M. Ziegler [27]. The diameter of a pointed polyhedron is the maximum, over all pairs of its vertices, of the minimum number of edges in a path joining them. A still open conjecture related to the empirical fact above is the (bounded) Hirsch conjecture: is the diameter of every bounded polyhedron in d -space with n facets bounded by $n - d$? The best known bound is considerably worse, although subexponential: G. Kalai and D. J. Kleitman [14] showed that it is at most $n^{\log d+2}$. For many simplex algorithms used in practice, although their performance in practice is overwhelmingly excellent, there are families of linear programming problems where the number of steps required grows exponentially in the dimensions of the problems. And finally, Kalai as well as J. Matousek, M. Sharir, and E. Welzl have developed randomized pivot rules whose expected number of steps on any linear programming problem on such a polyhedron is at most $\exp(K\sqrt{d\log n})$ for some constant K (see Kalai [13]).

These results do not explain the overwhelmingly effective performance of the simplex method. A number of authors considered the expected behavior of a particular simplex variant on a random linear programming problem drawn from some probability distribution on problems of given dimensions. For example, one such result, independently proved by I. Adler and N. Megiddo, by I. Adler, R. Karp, and R. Shamir, and by the reviewer, gives a bound of $O([\min\{d, n - d\}]^2)$ expected steps to prove infeasibility, establish unboundedness, or generate an optimal solution [1]. A good source for such results is the book [6] of K.-H. Borgwardt, who first established general polynomial-time results. All these theorems have a common drawback: the probability distributions used are highly suspect and are not believed to accurately represent linear programming problems that arise in practice. A very attractive result of D. A. Spielman and S.-H. Teng [23] addresses this difficulty and provides a beautiful interpolation between worst-case and average-case results: they give a simplex variant which, for any linear programming problem, requires an expected polynomial number of steps to solve a random problem with data drawn from a normal distribution centered at those of the nominal problem. The bound is a polynomial in the dimensions of the problem and in $1/\sigma$, where σ represents the standard deviation of the data. Thus the distribution can be highly concentrated about any particular instance. Such a result is called a *smoothed* analysis; see [24].

The shortcomings of the simplex method (at least theoretically) led to the search for other approaches to linear programming that had polynomial bounds on their behavior in the worst case. The first such method was a novel application of the ellipsoid method for convex programming problems introduced by D. B. Yudin and A. S. Nemirovski and independently by N. Z. Shor, developed and analyzed by L. G. Khachiyan [16, 17]. (Incidentally, we note that the field of computational complexity, a fundamental part of theoretical computer science, was largely motivated by the search for polynomial-time algorithms (often based on linear programming ideas) for certain combinatorial optimization problems; see the reminiscences of

J. Edmonds [9].) A survey of the ellipsoid method, with references to these papers and other related works, can be found in [5]. The algorithm, because of its original motivation, ignores almost completely the linear structure of the problem. It generates a sequence of ellipsoids with shrinking volumes, each guaranteed to contain all feasible solutions or all feasible solutions with objective value at least as good as some known solution. When applied to a linear programming problem with rational data, it either proves infeasibility, demonstrates unboundedness, or generates an optimal solution in a number of arithmetic operations bounded by a polynomial in the encoding length (bit length of all the data) of the problem. The polynomiality is related to the following geometric fact: if an ellipsoid in d -space is cut in half by a hyperplane through its center, each half-ellipsoid can be enclosed in a new ellipsoid whose volume is smaller than the original by a factor at most $\exp(-1/[2d + 2])$. Moreover, the new ellipsoid can be represented by parameters that are easily computed from those for the original ellipsoid and those describing the hyperplane. (This easy representability of an ellipsoid is important in the algorithm, distinguishing it from a general convex body which has the property that any hyperplane through its center of gravity divides it into two bodies, each of which has volume at most $1 - 1/e$ times that of the original body.)

The ellipsoid method turned out to be of no value at all for solving practical linear programming problems, although it remains a very valuable theoretical tool for analyzing the complexity of optimization problems, particularly those arising in combinatorics. See the excellent book of M. Grötschel, L. Lovász, and A. Schrijver [11] for these developments.

The next advance in linear programming algorithms was significant indeed. N. K. Karmarkar announced his projective algorithm for linear programming [15], together with a theoretical polynomial-time bound and a claim of practical superiority over the simplex method. The former was of great interest: the method used ideas of nonlinear programming like the ellipsoid algorithm, it performed a projective transformation at every step and then moved in a projected steepest-descent direction, and it employed a nonlinear potential function to measure progress. The sequence of points generated lay in the interior of the feasible region rather than moving from vertex to vertex, and although its theoretical worst-case time bound was hardly superior to that of the ellipsoid method, it appeared to be computationally viable and potentially competitive with the simplex method. The claim of superiority proved more troublesome, but with subsequent development, what are now called interior-point or barrier methods have proved highly efficient, particularly primal-dual variants. The new methods also provided a great spur to development of the simplex method, and the two approaches have vied ever since, with huge improvements to both and the comparison perhaps more dependent on coding and machine architecture than on the intrinsic algorithms. An excellent survey of the power of these algorithms, the size of the problems that can be solved, and the computational speedups due to both machines and algorithms from the late 1980s to about 2002, as well as a history of computational work on linear programming algorithms, appears in the paper of R. E. Bixby [4]. Both computer hardware improvements and algorithmic enhancements seem to have decreased times by three orders of magnitude, leading to an overall improvement of around a million-fold.

While interior-point methods were less flexible than the simplex method in starting from a good initial solution, crucial in solving integer programming problems, their significance for convex programming was immense. Y. E. Nesterov and

A. S. Nemirovski analyzed the reasons that these new methods were so successful for linear and quadratic programming and developed the notion of a *self-concordant barrier* for a convex set, which allows the derivation of an efficient barrier method to optimize a linear objective function over that set. This led to a huge expansion of the domain of interior-point methods; indeed, Nesterov and Nemirovski proved the existence of an efficient self-concordant barrier for *any* convex set in \mathfrak{R}^n , although in general it is not computable in any reasonable sense. They also developed a calculus of self-concordant barriers, and hence constructed computable barriers for many classes of feasible regions arising in practice. Two particular classes are the intersections of affine subspaces with the cone of positive semidefinite matrices (leading to semidefinite programming), and the intersections of affine subspaces with (products of) the second-order, Lorentz, or “ice-cream” cone (leading to second-order cone programming). These two classes have wide applicability and connections with symmetric cones [10]; indeed, the self-concordant barriers for these cones are the logarithms of the characteristic functions of the cones.

Second-order cone programming is relevant in robust optimization, in financial modelling, and in certain engineering design problems. Semidefinite programming has applications in eigenvalue optimization, in engineering design, in obtaining tight bounds in hard combinatorial optimization problems, and in global optimization of polynomials, where it connects with the *Positivstellensatz* of Schmüdgen and Putinar and with moment problems. The most comprehensive work on general interior-point methods and self-concordant barriers is the book of Nesterov and Nemirovski [20]; see also the readable monograph of J. Renegar [21], the book of A. Ben-Tal and Nemirovski [3] with theory and engineering applications, the book of J. B. Lasserre [18] on moment problems and polynomial optimization applications, and the survey articles [25, 19].

It is clear that optimization has come a long way since the early days of linear programming and remains a vibrant and exciting field with vast applicability and intriguing theory. But it remains highly valuable to examine its origins and early masters. The present volume, edited by his co-author and colleague of many years R. W. Cottle, collects some of the fundamental papers of George Dantzig from the 1940s through the early 1990s. It makes a strong claim that Dantzig was not only (as is widely acknowledged) the father of linear programming, but indeed of all of mathematical programming. Included are his basic papers on the simplex method; on the applications of linear programming to combinatorial optimization, including a landmark paper solving a 49-city traveling-salesman problem in 1954; on techniques for solving large-scale problems; on the linear complementarity problem; and on linear programming under uncertainty, an interest of Dantzig from his early paper in 1955 to his papers in the 1990s, and a challenge today. The book also includes Dantzig’s first two papers in mathematical statistics. The twenty-four chapters containing Dantzig’s papers are preceded by a preface by the editor, explaining the choice of papers included and providing a synopsis of Dantzig’s career and honors. Each part of the book, devoted to a collection of papers with a particular theme, has a short introduction to the papers therein as well as some associated history to put them in context. The papers themselves are rendered verbatim, but with helpful editor’s notes correcting typos, identifying the context of some remarks, and clarifying the text. These notes are collected, as well as a (long!) list of all of Dantzig’s publications, at the end of the book.

Sadly, George Dantzig (and Leonid Khachiyan) died in 2005. This book provides a lasting memorial to Dantzig's wide-ranging and seminal contributions to optimization.

The book can be highly recommended to all those in mathematical programming who are interested in the history of the field and wish to acquaint themselves (and be impressed) with the scope of its foundational papers. An excellent complement to the technical material here is a small volume on the history of mathematical programming, with articles by Dantzig [8] on linear programming and by M. Balinski [2] on many aspects of the early history, including the 1975 Nobel Prize in Economics for contributions to the theory of optimum allocation of resources, which was awarded to Kantorovich and Koopmans, but, inexplicably, disappointingly, and outrageously, not to Dantzig. Schrijver [22] provides an outstanding scholarly history of the field of combinatorial optimization. The reviewer's paper [26] describes developments in linear programming up to around 1999. And Dantzig's classic text [7] from 1963 describes his viewpoint on linear programming and its extensions at that time; it also includes much historical discussion.

REFERENCES

- [1] I. Adler, N. Megiddo, and M. J. Todd. New results on the average behavior of simplex algorithms. *Bulletin of the American Mathematical Society*, 11:378–382, 1984. MR752803 (85h:90074)
- [2] M. L. Balinski. Mathematical programming: Journal, society, recollections. In *History of Mathematical Programming*, (J. K. Lenstra, A. H. G. Rinnooy Kan, and A. Schrijver, editors), Elsevier Science Publishers, 1991, pp. 5–18.
- [3] A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. MPS-SIAM Series on Optimization. SIAM, Philadelphia, 2001. MR1857264 (2003b:90002)
- [4] R. Bixby. Solving real-world linear programs: A decade and more of progress. *Operations Research*, 50:3–15, 2002. MR1885204
- [5] R. Bland, D. Goldfarb, and M. Todd. The ellipsoid method: A survey. *Operations Research*, 29:1039–1091, 1981. MR641676 (83e:90081)
- [6] K. H. Borgwardt. *The Simplex Method: A Probabilistic Analysis*. Springer-Verlag, Berlin, 1986. MR868467 (88k:90110)
- [7] G. B. Dantzig. *Linear Programming and Extensions*. Princeton University Press, Princeton, NJ, 1963. MR0201189 (34:1073)
- [8] G. B. Dantzig. Linear programming. In *History of Mathematical Programming* (J. K. Lenstra, A. H. G. Rinnooy Kan, and A. Schrijver, editors), Elsevier Science Publishers, 1991, pp. 19–31. MR1183952 (94b:01030)
- [9] J. Edmonds. A glimpse of heaven. In *History of Mathematical Programming* (J. K. Lenstra, A. H. G. Rinnooy Kan, and A. Schrijver, editors), Elsevier Science Publishers, 1991, pp. 32–54. MR1183952 (94b:01030)
- [10] J. Faraut and A. Koranyi. *Analysis on Symmetric Cones*. Oxford University Press, Oxford, 1994. MR1446489 (98g:17031)
- [11] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*. Springer-Verlag, Berlin, 1988. MR936633 (89m:90135)
- [12] B. Grünbaum. *Convex Polytopes*. Wiley, New York, 1967. MR0226496 (37:2085)
- [13] G. Kalai. Linear programming, the simplex algorithm and simple polytopes. *Mathematical Programming*, 79:217–233, 1997. MR1464768 (98c:90065)
- [14] G. Kalai and D. J. Kleitman. A quasi-polynomial bound for the diameter of graphs of polyhedra. *Bulletin of the American Mathematical Society*, 24:315–316, 1992. MR1130448 (92h:52017)
- [15] N. K. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984. MR779900 (86i:90072)

- [16] L. G. Khachiyan. A polynomial algorithm in linear programming (in Russian). *Doklady Akademiia Nauk SSSR*, 224:1093–1096, 1979. English Translation: *Soviet Mathematics Doklady*, 20:191–194, 1979. MR522052 (80g:90071)
- [17] L. G. Khachiyan. Polynomial algorithms in linear programming (in Russian). *Zh. Vychisl. Mat. i Mat. Fiz.* 20:51–68, 1980. English Translation: *U.S.S.R. Computational Mathematics and Mathematical Physics*, 20:53–72, 1980. MR564776 (81j:90079)
- [18] J. Lasserre. *Moments, Positive Polynomials and Their Applications*. Imperial College Press, London, 2010.
- [19] A. S. Nemirovski and M. J. Todd. Interior-point methods for optimization. *Acta Numerica*, 17:191–234, 2008. MR2436012 (2009j:90141)
- [20] Y. E. Nesterov and A. S. Nemirovski. *Interior Point Polynomial Methods in Convex Programming: Theory and Algorithms*. SIAM Publications. SIAM, Philadelphia, USA, 1993.
- [21] J. Renegar. *A Mathematical View of Interior-Point Methods in Convex Optimization*. SIAM, Philadelphia, USA, 2001. MR1857706 (2002g:90002)
- [22] A. Schrijver. On the history of combinatorial optimization (til 1960). In *Handbook of Discrete Optimization* (K. Aardal, G. Nemhauser, and R. Weismantel, editors), Elsevier Science Publishers, 2005, pp. 1–68.
- [23] D. Spielman and S.-H. Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *Journal of the ACM*, 51:385–463, 2004. MR2145860 (2006f:90029)
- [24] D. Spielman and S.-H. Teng. Smoothed analysis: An attempt to explain the behavior of algorithms in practice. *Communications of the ACM*, 52:76–84, 2009.
- [25] M. J. Todd. Semidefinite optimization. *Acta Numerica*, 10:515–560, 2001. MR2009698 (2004g:90004)
- [26] M. J. Todd. The many facets of linear programming. *Mathematical Programming*, 91:417–436, 2002. MR1888985
- [27] G. Ziegler. *Lectures on Polytopes*. Springer-Verlag, Berlin, 1995. MR1311028 (96a:52011)

MICHAEL J. TODD

CORNELL UNIVERSITY

E-mail address: `mjt7@cornell.edu`