McMaster University

Advanced Optimization Laboratory

Title:

Pivot versus Interior Point Methods:
Pros and Cons

Author:

Tibor Illés and Tamás Terlaky

AdvOl-Report No. 2000/4

June 2000, Hamilton, Ontario, Canada

Pivot versus Interior Point Methods: Pros and Cons

Tibor Illés[1]
Department of Operations Research
Eötvös University, Budapest, Hungary

Tamás Terlaky[2]
Department of Computing and Software
McMaster University, Hamilton, ON, Canada

Abstract: Linear optimization (LO) is the fundamental problem of mathematical optimization. It admits an enormous number of applications in economics, engineering, basic sciences and many other fields. The three most significant classes of algorithms for solving LO are: Pivot, Ellipsoid and Interior Point Methods. Because Ellipsoid Methods are not efficient in practice we will concentrate on the computationally successful Simplex and Primal-Dual Interior Point Methods only, and summarize the Pros and Cons of these algorithm classes.

1 The LO problem.

In plain English one can say that a linear optimization problem consists of optimizing, i.e. minimizing or maximizing, a *linear function* over a certain domain. The domain is given by a set of *linear constraints*. The constraints can be either equalities or inequalities. For the first sight, LO problems have a quite simple structure. Only linear functions are involved, however not only a set of linear equations has to be solved, but our task is made more difficult in two ways. Inequalities are involved as well, and an optimization component "find a solution that has the best possible value of the objective function" is present. Before discussing mathematical properties and entering our main theme –pros and cons of the major algorithms– let us first devote some paragraphs to the history of LO. The two major components: optimizing a function and solving a linear system can be traced back for centuries.

Optimization is a natural activity. Optimization elements can be found in the ancient Greek mathematics too. Lagrange considered the minimization of certain functions while a set of –possibly nonlinear– equality constraints are satisfied. Although Lagrange has not considered inequalities, he had laid the foundations of duality theory.

On the side of solving linear systems an epoch making step was made by Gauss by establishing his celebrated *Gauss elimination* algorithm which is still one of the main techniques to solve systems of linear equations. However, besides equations, LO involves inequalities as well. An intellectually appealing, but computationally inefficient algorithm to solve systems of linear inequalities was designed by Fourier and Motzkin [11]. Systems of linear inequalities were thoroughly studied by Farkas [5]. He developed the theory of alternative systems. Farkas' fundamental result is still one of the most frequently cited theorem in the LO literature, it is essentially equivalent to the Strong Duality Theorem of LO.

Linear optimization –an important discipline with ever growing number of applications– was born in the middle of the 20th century as a new turbulent research field. Although some specially structured LO problems were formulated and systematically studied earlier by Kantorovich [11], Dantzig [4] was the first who developed a concise theory: geometrical analysis, duality theory, library of typical practical models and a computationally efficient algorithm, the simplex method which is still the most efficient pivot method to solve LOs. Illuminating true stories about the birth

[1]Partially supported by the Hungarian National Research Fund OTKA No. T 019492.
[2]Supported by the grant # OPG 0048923 of the National Sciences and Engineering Research Council of Canada and the Hungarian National Research Fund OTKA No. T 019492..

of LO and the simplex method can be found in [10], including Dantzig's recall of those historic years. Because solving large scale LO problems is a highly computation intensive activity, from the 1950's the algorithmic developments on the area of LO are closely related to the advances in computer technology. As the speed and capacity of computers increases, new avenues open to solve larger and larger LO problems. In the 1950's, problems with hundreds of variables were considered to be large scale problems. Today everyone can solve problems of tens of thousands of variables on his own desktop PC in some minutes and LO problems with millions of variables are already solved. The reader is referred to [1, 2] for surveys of computational state of the art.

The practical efficiency of the simplex method could not hide the frustrating fact that some variants of the simplex method might require prohibitively long time to solve some LO problems. This was made explicit in the early seventies by Klee and Minty [9] as they showed that a variant of the simplex method requires exponentially many steps to solve an LO problem where the set of feasible solutions is a slightly perturbed hypercube. On the other hand, under a probabilistic model, it is proved that the average case behavior of some simplex methods is polynomial [3].

The Klee-Minty example stimulated the research for alternative algorithms for LO. It was an open problem for long time if LO problems are polynomially solvable or not. The positive result was presented by Khachian [8]. Khachian's ellipsoid method was able to solve LO problems in polynomial time. This major result reached the front page of New York Times, but it turned out soon that in practice the ellipsoid method is computationally not competitive with the simplex method. In practice the simplex method performs much better than its theoretical worst case bound suggests, while the practical behaviour of the ellipsoid method is close to its worst case bound.

The next break-through was reached by Karmarkar [7] in 1984. He presented an interior point method (IPM) whose worst case complexity bound was better than for the ellipsoid method. Karmarkar also claimed superior computational performance. Karmarkar's results and claims started the interior point revolution [12, 13, 14, 15]. By now, the superior theoretical properties of IPMs are obvious; highly efficient implementations are available; the implementations are at least competitive with the implementations of simplex methods and they typically have superior performance on large scale problems [1].

In this paper we briefly list and compare the advantages and disadvantages of the two computationally highly successful major algorithm classes for solving linear optimization problems. We concentrate on features of Simplex Methods and Primal-Dual Path-following Interior Point Methods.

2 The primal dual LO problems

2.1 Duality: The primal and dual LO problem in standard form can be defined as

$$\text{minimize}\,\{\,c^T x\;:\;Ax = b,\;x \geq 0\,\},$$
$$\text{maximize}\,\{\,b^T y\;:\;A^T y + s = c,\;s \geq 0\,\},$$

where $c, x, s \in \mathbb{R}^n,\; b, y \in \mathbb{R}^m$ are vectors, and $A \in \mathbb{R}^{m \times n}$ is a given matrix. The deep, fundamental result, the so-called *Strong Duality Theorem* of LO is as follows.

Theorem 2.1 (Strong Duality) *If both the primal and dual problems admit a feasible solution, then for any optimal solutions* $\bar{x}$ *and* $\bar{y}$ *we have* $c^T\bar{x} = b^T\bar{y}$. *Moreover, if either the primal or the* *dual problem has no feasible solution and the other has a feasible solution, then the feasible problem is unbounded.*

Duality theorems can be used in many different ways. They can be used to check optimality of solutions. If two solutions $\bar{x}$ and $(\bar{y}, \bar{s})$ are given, then we need to check if $\bar{x}$ is primal feasible, i.e. if $A\bar{x} = b$, $\bar{x} \geq 0$ hold; if $(\bar{y}, \bar{s})$ is dual feasible, i.e. if $A^T\bar{y} + \bar{s} = c$, $\bar{s} \geq 0$ holds and finally, if the two objective values are equal, i.e. if $c^T\bar{x} = b^T\bar{y}$, or equivalently, if

$$\bar{x}_j(c - A^T\bar{y})_j = \bar{x}_j\bar{s}_j = 0 \quad \forall \;\; j = 1, \cdots, n.$$

The last relation says that optimal solutions are complementary. A stronger result, characteristic for LO problems, is stating that a *strictly complementary optimal solution* always exists.

Theorem 2.2 (Goldman-Tucker Theorem) *If both the primal and dual problems admit a feasible solution, then there are primal and dual optimal solutions $\bar{x}$ and $(\bar{y}, \bar{s})$ with $\bar{x} + \bar{s} > 0$, i.e., there exist a* strictly complementary *optimal solution pair.*

These relations provide the fundamentals of different algorithmic concepts what are fleshed briefly in the next subsection.

2.2 Fundamentals of algorithms: The duality theorem says that in order to find optimal solutions for both the primal and dual problems, we need to solve the system

$$\begin{array}{rrll} & Ax & = b, & x \geq 0, \\ A^Ty + & & s = c, & s \geq 0, \\ & & x_js_j = 0, & j = 1, \cdots, n. \end{array}$$

Here the first line represents primal feasibility, the second line is dual feasibility and the last one is the complementarity condition which guarantees that the primal and dual objective values are equal:

$$c^Tx = b^Ty.$$

The specific algorithms for LO approach to solve the system of optimality conditions differently. Pivot algorithms, including all Simplex Methods use basic solutions and hence they always keep the primal and dual equality and the complementarity conditions satisfied while nonnegativity of the variables does not necessarily hold. Further, primal (dual) simplex methods require and preserve primal (dual) feasibility while working towards optimality.

Interior-point methods (IPMs) seek to approach the optimal solution through a sequence of points that are always strictly feasible. Such methods have been known for a long time [6] but, because they are more demanding for storage and reliable floating point arithmetic than simplex methods, they were not considered to be effective. Contrary to pivot algorithms, interior point methods –at least the most popular primal-dual ones– require and preserve both primal and dual feasibility while the complementarity condition is relaxed. In fact, the requirement is even stronger, not only $x, s \geq 0$ is required but $x, s > 0$, i.e. both the primal and dual solutions are strictly positive. IPMs follow the so-called *central path.* We may assume w.l.g. that $\text{rank}(A) = m$ and strictly feasible solutions exist. Then the parameterized nonlinear system

$$\begin{array}{rrll} & Ax & = b, & x \geq 0, \\ A^Ty + & & s = c, & s \geq 0, \\ & & x_js_j = \mu, & j = 1, \cdots, n \end{array}$$

has a unique solution $(x(\mu), y(\mu), s(\mu))$ for each $\mu > 0$. The primal central path is defined as $\{x(\mu) \, : \, \mu > 0\}$ while the dual central path is $\{(y(\mu), s(\mu)) \, : \, \mu > 0\}$. Primal-Dual Path-following IPMs use Newton steps to follow the central path that leads the iterative process to an optimal solution.

3 Simplex versus Interior Point Methods

In this section the major features, pros and cons of simplex and IPMs are listed. At each feature first the properties/results related to simplex methods, then in italic the same issues regarding IPMs are addressed. Here we take the freedom to consider only Simplex Methods from the class of pivot methods and Primal-Dual Path-Following IPMs. Most results listed below can be found in the books [12, 14, 15]. Due to space limitations the list is not complete. Further we note that some of the statements do not have strong mathematical foundation, some of them are based on empirical evidence.

Characteristics of the iterative sequence: Simplex methods generate a sequence of feasible basic solutions. Complementarity of the candidate primal-dual solutions holds by definition. This geometrically means jumping from vertex-to-vertex on the boundary of the feasible set. The iterates are primal (or dual) feasible, the primal (or dual) objective improves monotonically; dual (or primal) feasibility is reached at optimality.

IPMs generate a sequence of strictly feasible primal and dual solutions. The iterates are in the (relative) interior of the feasible sets. complementarity is reached only at optimality. The duality gap decreases monotonically.

The generated solution: Simplex methods generate an optimal basic solution. In case of degeneracy the found optimal basic solution is not strictly complementary. To find a strictly complementary solution from an optimal basis is not easy, it is just as difficult as solving the original problem.

IPMs produce an ϵ optimal solution, i.e. a feasible solution pair for which $c^T x - b^T y < \epsilon$. Starting from a sufficiently precise solution an exact strictly complementary solution can readily be identified. Moreover, an optimal basis can be found by using at most n additional pivots.

Complexity: No polynomial time version of the simplex method is known. Exponential examples exist for most variants.

Interior point methods enjoy polynomial time worst case complexity. So far the best known iteration bound is $\mathcal{O}(\sqrt{n}L)$ Newton steps which cost each $\mathcal{O}(n^{2.5})$ arithmetic operations resulting in an $\mathcal{O}(n^3 L)$ total complexity.[3] *The cost of identifying of an exact strictly complementary solution is $\mathcal{O}(n^3)$ arithmetic operations, while an optimal basis can be found by using at most n pivots.*

Initialization: LO problems are solved in two-phases because the simplex method requires a feasible basic solution to start with. Such a solution is readily available for the so-called first-phase problem and the optimal basic solution of this first phase problem provides a feasible basic solution for solving the original problem.

IPMs require a positive vector to start with. When one applies an infeasible IPM *then the Newton procedure makes sure that finally the equality constraints will be satisfied as well. The price of this simple direct approach is some weaker complexity result. The best approach is based on the so-called* self-dual embedding model *which can also be initialized by any positive vector. By marginal increase of the computational cost per iteration the most efficient feasible IPMs can be applied for the embedding model and this way the LO problem can be solved in one phase.*

Degeneracy: Simplex methods, both theoretically and in practice suffer from problems –especially cycling and stalling– arising from degeneracy of LO problems. Special techniques, such as

[3] Here L denotes the input length of the given LO problem.

lexicographic and least index rules, various heuristics are developed to circumvent these problems.

Degeneracy is not an issue in IPMs. Theoretical complexity results hold without any non-degeneracy assumption and, what is at least so important, the practical performance of IPMs is not affected by degeneracy.

Software: The state-of-the-art implementations of simplex methods, e.g., XPRESS-MP, CPLEX and OSL keep competing with IPM based software. The simplex method is very flexible, it allows the implementation of various heuristics to enhance practical performance.

Various highly efficient IPM based software were developed in the last years. Here we just mention some of the best known ones: BPMPD, MOSEK, HOPDM, CPLEX-Barrier, LOQO, PCx, LIPSOL.

Practical performance: The state-of-the-art implementations of simplex methods keep competing with IPM based software. In various areas still they are the method of choice. In practice simplex methods hardly need more than $2(m+n)$ pivots to solve an LO problem.

IPM software is especially efficient for solving very-large scale LO problems. When solving huge problems, possibly involving millions of variables, and solving highly degenerate problems IPMs outperform simplex method based codes. In practice IPMs need about 15-30 Newton steps to solve LO problems.

Note that all modern LO packages contain both simplex and IPM solvers.

Warm-start: Having the problem solved, the LO problem frequently need to be solved again with slightly changed data. Re-starting a simplex algorithm from the previous optimal solution, in most cases, allows quick solution of the modified problem.

Although many efforts, sometimes with promising results, were invested to design efficient warm-start procedures, IPM codes do not exhibit such efficiency in re-solving LO problems with perturbed data as simplex based software does.

Efficiency when solving integer problems: To date simplex methods are the clear winner in solving (mixed) integer LO problems. One of the reasons is that to generate cuts a basic solution is needed but, more importantly, the recently implemented dual simplex methods are incredibly efficient in branch-and-bound schemes.

As mentioned earlier, a basic solution can always be obtained by using IPM codes when proper basis identification schemes are implemented. Therefore cutting plane methods are applicable by using IPMs as well. The lack of efficient warm-start procedures make IPMs inefficient for solving integer LO problems.

Sensitivity Analysis: For long time it was believed that a basic solution is needed for sensitivity analysis. In-spite of this belief, as it was demonstrated by various authors, classical sensitivity analysis gives false sensitivity information when the LO problem is degenerate.

Due to the birth of IPMs, the theory of sensitivity and parametric analysis was revisited. It turned out that sensitivity analysis can be made by having any *optimal solution pair and, in case of degeneracy, an optimal basis does not contain any useful extra information regarding this question. To the contrary, a maximally complementary solution given by IPMs is more suitable for making correct sensitivity analysis.*

Generalizations: Simplex methods provide a solid base for various nonlinear optimization algorithms, such as complementary pivot algorithms to solve linear complementary problems, (generalized) reduced gradient methods and decomposition methods to solve semi-infinite optimization problems.

All the areas where simplex-like methods were adapted, IPMs were applied as well. On the top of those, IPMs provide most efficient algorithms to solve conic-linear optimization problems. First of all semi-definite and second-order cone optimization problems are formed and *solved with success by using IPMs. Moreover, semi-definite optimization allows powerful approximation schemes for various hard combinatorial optimization problems. The obtained approximation results are significantly better than those obtained by using linear relaxations.*

As seen from our discussions, there is no clear winner in the race to solve LO problems in practice. The research is going on to find better algorithms both considering the theoretical worst case complexity and practical performance.

References

[1] E.D. Andersen, J. Gondzio, Cs. Mészáros, and X. Xu. Implementation of interior point methods for large scale linear programming. In T. Terlaky, editor, *Interior Point Methods of Mathematical Programming*, pages 189–252. Kluwer A.P., Dordrecht, The Netherlands, 1996.

[2] R.E. Bixby. Progress in linear programming. *ORSA J. on Computing*, 6(1):15–22, 1994.

[3] K.H. Borgwardt. *The Simplex Method: A Probabilistic Analysis.* Springer Verlag, Berlin, 1987.

[4] G.B. Dantzig. *Linear Programming and Extensions.* Princeton University Press, Princeton, New Jersey, 1963.

[5] J. Farkas. Theorie der einfachen Ungleichungen. *J. Reine und Angewandte Mathematik*, 124: 1–27, 1902.

[6] A.V. Fiacco and G.P. McCormick. *Nonlinear programming: sequential unconstrained minimization techniques.* John Wiley & Sons, New York, 1968. (Reprint: Volume 4 of *SIAM Classics in Applied Mathematics*, SIAM Publications, Philadelphia, USA, 1990).

[7] N.K. Karmarkar. A new polynomial–time algorithm for linear programming. *Combinatorica*, 4: 373–395, 1984.

[8] L.G. Khacijan. A polynomial time algorithm in linear programming. *Soviet Mathematics Doklady*, 20:191–194, 1979.

[9] V. Klee and G.J. Minty. How good is the simplex algorithm? In O. Shisha, editor, *Inequalities III.* Academic Press, New York, 1972.

[10] J.K. Lenstra, A.H.G Rinnooy-Kahn and A. Schrijver, editors. *History of mathematical programming. A collection of personal reminiscences.* CWI, North–Holland, The Netherlands, 1991.

[11] A. Prékopa. On the development of optimization. *Amer. Math. Monthly*, 87:527–542, 1980.

[12] C. Roos, T. Terlaky, and J.-Ph. Vial. *Theory and Algorithms for Linear Optimization: An Interior Point Approach.* John Wiley & Sons, New York, 1997.

[13] M.H. Wright. The interior-point revolution in constrained optimization. Numerical Analysis Manuscript 98–4–09, AT & T Bell Lab's, Murray Hill, USA, 1998.

[14] S. Wright. *Primal-Dual Interior-Point Methods.* SIAM, Philadelphia, 1996.

[15] Y. Ye. *Interior Point Algorithms.* John Wiley & Sons, New York, 1997.