WMO (1976), Chang (1977) and Haltiner and Williams (1980).

Despite the progress that has been made, it appears likely that there is still a long way to go before the ideal numerical method is found which integrates the governing equations and gives clearly maximum accuracy for a given computational cost.

References

Bates, J.R. and A. McDonald, 1982: Monthly Weather Review, 110, No. 12 (in press).

Chang, J. (Editor), 1977: Methods in Computational Physics, Vol. 17, Academic Press.

Charney, J., R. Fjørtoft and J. von Neumann, 1950: Tellus, 2, 237-254.

Courant, R., K. Friedrichs and H. Lewy, 1928: Math. Ann., 100, 32-74.

Haltiner, J. and R.T. Williams, 1980: Numerical Prediction and Dynamic Meteorology, John Wiley and Sons.

Kwizak, M. and A. Robert, 1971: Monthly Weather Review, 99, 32-36.

Marchuk, G.I., 1974: Numerical Methods in Weather Prediction, Academic Press.

WMO, 1976: GARP Publications Series, No. 17, Vols. 1 and 2, World Meteorological Organization, Geneva.

Meteorological Service,
Glasnevin Hill,
Dublin 9.

RECENT DEVELOPMENTS IN LINEAR PROGRAMMING

Fergus J. Gaines

Let A be an mxn matrix, let $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$. The basic problem in linear programming is to find, for $x \in \mathbb{R}^n$,

$$\max c^t x, \text{ subject to } Ax \leq b,\ x \geq 0 \qquad (1)$$

For vectors, $x \leq y$ means $x_i \leq y_i$ for all i; $x < y$ means $x_i < y_i$ for all i.)

The standard way of solving this problem is to use the celebrated *simplex method* of G. Dantzig [1]. The idea is to note that the *feasible* solutions of (1), i.e. the $x \in \mathbb{R}^n$ with $Ax \leq b$, $x \geq 0$, form a convex polytope K in $\mathbb{R}^n$. The vertices of K are those feasible x with either x = 0 or such that the positive components of x correspond to linearly independent columns of A. The typical step in the simplex algorithm proceeds from vertex $x^{(k)}$ to a vertex $x^{(k+1)}$ so that $c^t x^{(k+1)} \geq c^t x^{(k)}$. Since max $c^t x$ is attained at a vertex of K, the algorithm eventually gives the answer.

This algorithm is arguably the most widely used algorithm of the present day and it is probably safe to say that most of those who use it do not understand it, whereas most of those capable of understanding it never use it. Its popularity is probably the reason for the widespread, if in many cases inaccurate, coverage in the newspapers given to the discovery in 1979 of a new algorithm for solving (1), the work of a Soviet "unknown" L.G. Khachiyan [2]. (One American newspaper reported bitterly (but incorrectly) that a Soviet mathematician had solved the "travelling salesman problem", despite the fact that the U.S.S.R has no travelling salesmen!)

The immediate reason why Khachiyan's algorithm is important is because it is *in theory* more computationally efficient than the simplex method. One of the noteworthy features of the simplex algorithm (and its variants) is that it is very efficient in all practical cases, i.e. it uses very little machine time. Empirical data show that the number of operations (+,x, etc.) in a typical application is $O(mn^3)$. However, Klee and Minty [3] have produced an example with $m=2n$ where the simplex method requires more than 2^n steps. In contrast, Khachiyan's algorithm is "polynomially bounded" in all cases, but it has serious drawbacks (see below).

But why does the simplex method work so well in practice? In a recent, highly significant paper, [4], Steve Smale has given a very satisfactory explanation. We discuss Smale's result below.

Khachiyan's Algorithm

Since Khachiyan's paper contains no proofs we follow the presentation in [5]. We note that the *linear programme* (LP) (1) can be reduced to the problem of solving a system of linear inequalities. We see this as follows. With LP (1) we can associate the *dual* LP, which is to find, for $y \in \mathbb{R}^m$

$$\min b^t y, \text{ subject to } A^t y \geq c. \quad y \geq 0 \qquad (2)$$

The Duality Theorem says (1) has an optimal solution if and only if (2) has, and in the event, $\max c^t x = \min b^t y$. Thus (1) has a finite optimum if and only if the system of inequalities

$$Ax \leq b, \quad x \geq 0, \quad A^t y \geq c, \quad y \geq 0, \; c^t x \geq b^t y \qquad (3)$$

has a solution. If (x,y) is a solution of (3) then x is an optimal solution of (1). The inequalities (3) can be rewritten

$$Mz \leq d, \qquad z \geq 0$$

where

$$M = \begin{bmatrix} A & 0 \\ 0 & -A^t \\ -c^t & b^t \end{bmatrix} \qquad z = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \qquad \text{and } d = \begin{bmatrix} b \\ -c \\ 0 \end{bmatrix}$$

So (changing notation) we need only solve the problem: find $x \in \mathbb{R}^n$ with $Ax \geq b$, $x \geq 0$.

We describe the algorithm for the problem:

$$\text{find } x \in \mathbb{R}^n \text{ with } Ax < b, \; x \geq 0, \text{ where A and B have integer entries.} \qquad (4)$$

(The case $Ax \leq b$ can be reduced to this case)

At first sight, the restriction to integer entries does not appear significant as the practical implementation of any algorithm can only involve finite decimals which are equivalent to integers. But this seems to be the root cause of the bad behaviour of the algorithm in practice.

The algorithm determines a sequence $x^{(k)} \in \mathbb{R}^n$ and a sequence of ellipsoids $E^{(k)}$ in $\mathbb{R}^n$ with centre $x^{(k)}$ and

$$\text{vol}(E^{(k+1)}) < \text{vol}(E^{(k)})$$

If L is the length of the binary encoding of (4) the algorithm either gives, for some $k < 4(n+1)^2L$, an $x^{(k)}$ which is a solution of (4) or, if a solution cannot be found for such k, it shows that no solution exists.

If B_k is a positive definite symmetric matrix then

$$E^{(k)} = \{x \in \mathbb{R}^n : (x-x^{(k)})^t B_k^{-1} (x-x^{(k)}) \leq 1\}$$

is an ellipsoid with centre $x^{(k)}$. The steps in the algorithm are:

1. Set $x^{(0)} = 0$, $B^{(0)} = 2^2 VI$.
2. If $x^{(k)}$ is a solution to (4), terminate. If $k < 4(n+1)^2L$ go to 3. Otherwise terminate, concluding (4) has no solution.
3. Choose one of the inequalities in (4) not satisfied by $x^{(k)}$, say $a_i^t x^{(k)} \geq b_i$ (a_i^t is the i th row of A).

Let

$$x^{(k+1)} = x^{(k)} - (1/(n+1))B^{(k)}a_i/(a_i{}^tB^{(k)}a_i)^{\frac{1}{2}}$$

and

$$B^{(k+1)} = (n^2/n^2-1)[B^{(k)} - (2/n+1)(B^{(k)}a_i)(B^{(k)}a_i)^t/(a_i{}^tB^{(k)}a_i)]$$

Go to step 2 with k+1 in place of k.

The ellipsoid $E^{(k+1)}$ contains the semi-ellipsoid

$$E^{(k)} \cap \{x \in \mathbb{R}^n : a_i{}^t(x-x^{(k)}) \leq 0\}$$

Also

$$\mathrm{vol}(E^{(k+1)}) = c(n)\mathrm{vol}(E^{(k)})$$

where

$$c(n)^{2n-2} = \tfrac{1}{2}$$

The ellipsoid algorithm in the worst case is $O(n^3(m+n)L)$ in contrast to the exponential behaviour of the Klee-Minty example. However, the ellipsoid algorithm behaves very badly in practice. As Dantzig points out (cf. [6]) a typical economic planning problem which takes half an hour machine time for the simplex method to solve, would take the ellipsoid algorithm fifty million years! Traub and Wozniakowski [6] give an explanation for the poor performance of Khachiyan's algorithm. They show that for the real number computational model (i.e. $\mathbb{R}$ with exact arithmetic and unit "cost" for each operation) the ellipsoid algorithm in the worst case is not polynomially bounded.

Despite its failure to oust the simplex method, the ellipsoid algorithm appears to have a future in the solution of combinatorial optimization problems other than linear programming. The paper [7] of Grötschel, Lovasz and Schrijver deals with this topic.

<u>Smale's Theorem</u>

Dantzig ([1], p.160) conjectured that for a randomly chosen LP, with fixed number of constraints m, the number of operations in the simplex method grows in proportion to n. Smale [4] not only proved this result but improved on it considerably.

The first problem is to define the average number of steps in the simplex method for a LP. We get a probability measure μ on the unit sphere S^{p-1} in $\mathbb{R}^p$, by normalizing the standard uniform (Lebesgue) measure. The points of S^{p-1} correspond to the rays of $\mathbb{R}^p$. If X is a set of rays in $\mathbb{R}^p$, we define the *spherical measure* of X by $\nu(X) = \mu(X \cap S^{p-1})$. Let A,b,c be as in (1). Then $q = (c,-b) \in \mathbb{R}^N$, where $N=m+n$. Let $\sigma(A,q)$ be the number of steps required to solve (1) by the simplex method. Since $\sigma(A,\lambda q) = \sigma(A,q)$ for $\lambda > 0$, we identify q with a ray in $\mathbb{R}^N$. The average number of steps required to solve (1), with A fixed, is

$$\rho_A = \int_{q \in S^{p-1}} \sigma(A,q)\, d\mu$$

Now identify the space $\mathcal{A}$ of all real mxn matrices with $\mathbb{R}^{mn}$. Since $\sigma(\lambda A,q) = \sigma(A,q)$ for $\lambda > 0$ we identify A with an element of $\mathcal{A}_1$, the set of rays of $\mathcal{A}$. Put a spherical measure ν on $\mathcal{A}_1$. Then the average number of steps required to solve (1) is

$$\rho(m,n) = \int_{A \in \mathcal{A}_1} \rho_A\, d\nu$$

We now have Smale's result.

<u>Theorem</u> Let p be a positive integer. Then depending on p and m, there is a positive constant c_m such that for all n

$$\rho(m,n) \leq c_m n^{1/p}$$

The case p=1 is Dantzig's conjecture.

The proof of the theorem is not easy. Smale considers a version of the simplex method, Lemke's algorithm, applied to the *linear complementarity problem* (LCP): given an NxN real matrix M and $q \in \mathbb{R}^N$, find $w, z \in \mathbb{R}^N_+$, the positive orthant, so that $w^t z = 0$ and $w - Mz = q$. The primal-dual problem (3) is a special case of the LCP. Next he defines a mapping Φ_M on $\mathbb{R}^N$ so that the LCP becomes: find $x \in \mathbb{R}^N$ so that $\Phi_M(x) = q$. If $q_0 = (1, \ldots, 1)^t \in \mathbb{R}^N$, the inverse image of the line segment qq_0, $\Phi_M^{-1}(qq_0)$, is a piecewise linear curve γ in $\mathbb{R}^N$. If γ_0 is the component of γ containing q_0 then Lemke's algorithm can be viewed geometrically as "following" γ_0. A pivot of the algorithm corresponds to the intersection of γ_0 with a *facet* (a facet is the intersection of a hyperplane with an orthant Q_S; for

$$S \in \{1,2, \ldots, N\},\ Q_S = \{x \in \mathbb{R}^N : x_i \geq 0,\ i \in S,\ x_j \leq 0,\ j \notin S\}).$$

There are three main steps in the proof of the theorem. Firstly he derives a formula for ρ_A in terms of the spherical volume of certain cones. Then he derives an estimate for ρ_A. Finally he gets a simplified version of this estimate, when m is fixed and n is large, which gives the result.

The problem of determining the average speed of the simplex method as a function of both m and n still remains. In his Dublin lecture (September 1982) Smale said he felt that his general estimate for ρ_A might be used to solve this problem. However, the basic difficulty to be overcome is that of determining volumes of cones.

References

1. G. Dantzig. Linear Programming and Extensions, Princeton University Press, Princeton, N.J. (1963).

2. L.G. Khachiyan. A Polynomial Algorithm in Linear Programming (Russian), Doklady Akad. Nauk. S.S.S.R., 244 (1979), 1093-1096 (English translation in Soviet Mathematics Doklady, 20 (1979), 191-194).
3. V. Klee and G. Minty, How Good is the Simplex Algorithm? in O. Shisha, Ed.: Inequalities III, Academic Press, New York (1972), 159-175.
4. S. Smale, On the Average Speed of the Simplex Method of Linear Programming (to appear).
5. B. Aspvall and R.E. Stone, Khachiyan's Linear Programming Algorithm, J. Algorithms, 1 (1980), 1-13.
6. J. Traub and H. Wozniakowski, Complexity of Linear Programming, Operations Research Letters 1 (1982), 59-62.
7. M. Grötschel, L. Lovasz and A. Schrijver, The Ellipsoid Method and its Consequences in Combinatorial Optimization, Combinatorica 1 (1981), 169-197.

Mathematics Department,
University College Dublin.