Journal of Computational Mathematics, Vol.8, No.1, 1990, 16–22.

A METHOD OF FINDING A STRICTLY FEASIBLE SOLUTION FOR LINEAR CONSTRAINTS*

Wei Zi-luan
(*Computing Center, Academia Sinica, Beijing, China*)

Abstract

This paper presents a method of finding a strictly feasible solution for linear constraints. We prove, under certain assumptions, that the method is convergent in a finite number of iterations, and give the sufficient and necessary conditions for the infeasiblity of the problem. Actually, it can be considered as a constructive proof for the Farkas lemma.

§1. Introduction

In this paper we consider the following problem: to find a vector $x^{(0)} > 0$ which satisfies the linear constraints

$$Ax = b, \quad x \geq 0 \tag{1.1}$$

where A is an $m \times n$ real matrix with rank m, b is a real vector in R^m, and x is a real variable in R^n. A vector $x^{(0)}$ is called a strictly feasible solution if it satisfies (1.1) and all its components are positive.

This problem arises in solving the standard form of linear programming using an interior point method [7], [8], and minimizing the problem of a nonlinear objective function with linear constraints by means of barrier and penalty functions. Especially, a new polynomial-time algorithm for linear programming [4] was presented in recent years, It is a great improvement on complexity, and furthermore, is said to be 50 times faster than the simplex method for practical problem, Unfortunately, no further information on the test problems or experimental procedures was given. Therefore, it has arosed extensive attention and discussion. The idea of the new algorithm originated from the techniques of solving nonlinear programming problems. Obviously, the essential difference between the new algorithm and the simplex method is that it finds the optimal solution from the interior feasible direction of the constrained region. On the other hand, a similar result can also be deduced from a projected Newton barrier function [2] and the penalty function method [3]. As is well known, these methods all require a strictly feasible starting point for minimization, and generate a sequence of strictly feasible solution. So, how to find an initial strictly feasible solution for problem (1.1) in practice is an important problem.

This paper presents an efficient method of finding a strictly feasible solution for problem (1.1). In fact, the method can be introduced directly from the interior point method, and it leads to computational simplicity. In Section 2 we describe the algorithm, and show how to start it, when to stop it, and how to easily identify infeasibility. In Section 3, under certain assumptions, we prove its convergence to a strictly feasible solution in a finite number of iterations, and the sufficient and necessary conditions for the infeasibility.

* Received December 29, 1986.

§2. Feasibility and Algorithm

In this section, the necessary and sufficient conditions of the feasibility for problem (1.1) are briefly discussed, and the sufficient conditions of the existence of a strictly feasible solution are given, we give an algorithm for finding a strictly feasible solution in a finite number of iterations or indicating infeasible conditions for problem (1.1) in the case of nondegeneration, because the degeneration case is too complicated to be discussed here.

Concerning the feasibility, actually Farkas's theorem has shown the necessary and sufficient condition of feasibility for problem (1.1).

Lemma 2.1 (Farkas Theorem). *Suppose that $A \in R^{m\times n}, b \in R^m$. Then problem (1.1) is feasible if and only if for all the nonzero vectors $y \in R^m$ which satisfy $A^T y \geq 0$, the following inequality holds:*

$$b^T y > 0.$$

Obviously, the existence of the strictly feasible solution is not garanteed when problem (1.1) is feasible, Therefore, it is necessary to have a strong condition in order to ensure the existence of a strictly feasible solution.

Theorem 2.2 *Suppose that $rank(A) = m$, and that problem (1.1) is feasible and nondegenrate. Then there is a strictly feasible solution.*

A constructive proof of the theorem is given in Section 3.

Now we describe our algorithm. Assume that $x^{(0)} > 0$ is a given vector.

Algorithm A:

Let $k = 0$, and let an initial starting point $x^{(0)}$, be given

(1) Define

$$D_k = \mathrm{diag}(x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)}), \tag{2.1}$$

$$A_k = AD_k \tag{2.2}$$

and compute the residual vector

$$r^{(k)} = b - Ax^{(k)}. \tag{2.3}$$

(2) Compute vector

$$p^{(k)} = A_k^T (A_k A_k^T)^{-1} r^{(k)}. \tag{2.4}$$

If $p^{(k)} \leq 0$, and $b^T (A_k A_k^T)^{-1} r^{(k)} > 0$, stop; then problem (1.1) is infeasible. Otherwise, go to the next step.

(3) Choose the minimum component of $p^{(k)}$, and let

$$P_k = \min_i \{p_i^{(k)}\}. \tag{2.5}$$

If $-1 < P_k$, then

$$x^{(k+1)} = x^{(k)} + D_k p^{(k)}. \tag{2.6}$$

Thus, $x^{(k)}$ is a strictly feasible solution for problem (1.1), stop. Otherwise, go to next step.

(4) Select a suitable value of α_k, and compute

$$x^{(k+1)} = x^{(k)} + \alpha_k D_k p^{(k)} \tag{2.7}$$

such that $x^{(k+1)} > 0$, where $\alpha_k \in (0,1)$.

Let $k : k+1$, and return to step (1).

Actually, algorithm A can be deduced directly from the interior point method for linear programming [8]. This will be discussed briefly in Section 3 in order to prove the necessary conditions for the infeasibility.

It is clear that the main computational work at each iteration is from step (2), which ensures reduction of the Euclidean length of the residual vector.

§3. Finite Convergence

This section deals with the finite convergence of algorithm A, and the sufficient and necessary conditions for infeasibility of problem (1.1).

The main results are described as follows:

Theorem 3.1. *Suppose that problem (1.1) is feasible and nondegenerate, and that rank(A) = m. Then a strictly feasible solution can be obtained in a finite number of iterations by algorithm A.*

Theorem 3.2. *Suppose that the rank of matrix A is m. Then problem (1.1) is infeasible if and only if there exists an integer $k > 0$, such that*

$$p^{(k)} \le 0, \quad b^T (AD_k^2 A^T)^{-1} r^{(k)} > 0 \tag{3.1}$$

where $p^{(k)}$ is defined by (2.4).

The following two lemmas are introduced in order to prove the two theorems.

Lemma 3.3. *Suppose that the rank of matrix A is m, and that problem (1.1) is infeasible. Assume $y^{(k)} = (AD_k^2 A^T)^{-1} r$. Then there exists an integer $k > 0$, such that*

$$A^T y^{(k)} \le 0, \quad b^T y^{(k)} > 0 \tag{3.2}$$

where D_k and r are defined by (2.1) and (2.3), respectively.

Proof. Now Consider the linear programming LP_1 which is equivalent to problem (1.1):

$$LP_1: \quad \begin{array}{ll} \text{Min} & t \\ \text{s.t.} & Ax + rt = b \qquad (3.3) \\ & x, t \ge 0 \qquad (3.4) \end{array}$$

where $r = b - Ax^{(0)}$, and $x^{(0)} > 0$ is a given vector. Obviously, $x = x^{(0)}$, and $t = 1$ is a feasible solution of problem. Thus, the minimum value of porblem LP_1 is greater than zero since problem (1.1) is infeasible.

Assume that a sequence $\{x^{(l)}, t_l\}$ is generated by using interior point method to solve problem LP_1. Then

$$\begin{pmatrix} x^{(k+1)} \\ t_{k+1} \end{pmatrix} = \begin{pmatrix} x^{(k)} \\ t_k \end{pmatrix} + \delta_k \begin{pmatrix} D_k & \\ & t_k \end{pmatrix} C_P^{(k)} \tag{3.5}$$

where $\delta_k \in (0,1)$, and $C_p^{(k)}$ is a projected vector of the objective function on the null space of matrix (A, r), that is,

$$\begin{aligned} C_p^{(k)} &= \begin{pmatrix} D_k & \\ & t_k \end{pmatrix} \left[\begin{pmatrix} 0 \\ 1 \end{pmatrix} - \begin{pmatrix} A^T \\ r^T \end{pmatrix} (AD_k^2A^T + t_k^2 rr^T)^{-1} r t_k^2 \right] \\ &= \begin{pmatrix} -D_k A^T (AD_k^2A^T + t_k^2 rr^T)^{-1} r t_k^2 \\ t_k[1 - r^T(AD_k^2A^T + t_k^2 rr^T)^{-1} r t_k^2] \end{pmatrix}. \end{aligned} \tag{3.6}$$

By Sherman–Morrison–Woodbury formula[5], it is easy to show that

$$(AD_k^2A^T + t_k^2 rr^T)^{-1} r = \frac{(AD_k^2A^T)^{-1} r}{1 + t_k^2 r^T (AD_k^2A^T)^{-1} r}. \tag{3.7}$$

Substituting (3.7) into (3.6) gives

$$C_p^{(k)} = \frac{t_k}{1 + t_k^2 r^T (AD_k^2A^T)^{-1} r} \cdot \begin{pmatrix} -D_k A^T (AD_k^2A^T)^{-1} r\, t_k \\ 1 \end{pmatrix}. \tag{3.8}$$

Let

$$\bar{y}^{(k)} = \frac{(AD_k^2A^T)^{-1} r t_k^2}{1 + t_k^2 r^T (AD_k^2A^T)^{-1} r}, \tag{3.9}$$

$$\bar{\alpha} = \frac{\delta_k t_k}{1 + t_k^2 r^T (AD_k^2A^T)^{-1} r}. \tag{3.10}$$

Thus, it follows from (3.5), and (3.8)–(3.10) that

$$x^{(k+1)} = x^{(k)} + \delta_k D_k A^T \bar{y}^{(k)}, \tag{3.11}$$

$$t_{k+1} = (1 - \bar{\alpha}_k) t_k. \tag{3.12}$$

Because problem (1.1) is infeasible, there exists a $t^* > 0$ such that

$$\lim_{k \to \infty} t_k = t^*.$$

From (3.9), it is easy to show that for all integers $k > 0$,

$$r^T \bar{y}^{(k)} = \frac{t_k^2 r^T (AD_k^2A^T)^{-1} r}{1 + t_k^2 r^T (AD_k^2A^T)^{-1} r} < 1.$$

Evidently, $\bar{y}^{(k)}$ is an approximate solution for the dual problem of problem LP_1. It follows from the strong dual theorem of linear programming that there exists a sufficiently large integer $k_0 > 0$, such that for all $k > k_0$,

$$D_k A^T \bar{y}^{(k)} \leq 0, \quad b^T \bar{y}^{(k)} > 0.$$

From (3.9) and definition (2.1), it is clear that

$$\frac{t_k^2}{1 + t_k^2 r^T (AD_k^2A^T)^{-1} r} > 0$$

and

$$D_k > 0.$$

Thus, let

$$y^{(k)} = (AD_k^2A^T)^{-1}r.$$

Then, it is obvious that for all $k > k_0$,

$$A^Ty^{(k)} < 0, \qquad b^Ty^{(k)} > 0.$$

which proves the lemma.

It follows from (3.11) and (3.12) that the interior point method for solving problem LP_1 is equivalent to algorithm A. Because the new iterative points generated by the two methods are both updated by the same direction $p^{(k)} = (AD_k)^T(AD_k^2A^T)^{-1}r^{(k)}$, the same sequences $\{x^{(k)}\}$ can be generated by the two methods as long as we choose suitable α_k and δ_k.

Lemma 3.4. *Suppose that the rank of matrix A is m, and that problem (1.1) is feasible and nondegenerate. Assume that the sequence $\{x^{(k)}\}$ is generated by algorithm A. Then $E = \min_k\{E_k\} > 0$, where E_k is the Euclidean norm about $x^{(k)}$.*

Proof. Let $S = \{\begin{pmatrix} x \\ t \end{pmatrix} : Ax + rt = b, x \geq 0, t \geq 0, x \in R^n, t \in R^1\}$, where $r = b - Ax^{(0)}$, and $x^{(0)} > 0$ is a given vector. It is clear that $x = x^{(0)}, t = 1$ is a feasible point of S. So, S is a nonempty polyhedron. It is also easy to show that neither $x = 0, t = 0$ nor $x = 0, t > 0$ is in set S, it follows that there must exist a point $\bar{x}, \bar{t}$ in S such that

$$\bar{\rho} = \sqrt{\|\bar{x}\|_2^2 + \bar{t}^2} = \min_{\binom{x}{t} \in S} \left\{(\|x\|_2^2 + t^2)^{1/2}\right\}.$$

Hence, for all k,

$$\rho_k = \sqrt{\|x^{(k)}\|_2^2 + t_k^2} \geq \bar{\rho} > 0.$$

The minimum value of LP_1 defined by (3.3) and (3.4) is zero because of the feasibility and nondegeneracy of problem (1.1). Therefore, for a given sufficiently small $\varepsilon > 0$, there exists an integer $k_0 > 0$ such that for all $k > k_0, t_k < \varepsilon$. Thus, it is straightforward to establish

$$\|x^{(k)}\|_2 \geq \sqrt{\bar{\rho}^2 - t_k^2} > \bar{\rho} - \varepsilon.$$

Hence

$$E = \min_k\{E_k\} \geq \min\{\min_{k \leq k_0}\{E_k\}, \bar{\rho} - \varepsilon\} > 0$$

which proves the lemma.

Lemma 3.5. *Under the assumptions as in Lemma 3.4, there exists a large $M > 0$, such that for all k,*

$$\|D_kA^T(AD_k^2A^T)^{-1}r\|_2 < M.$$

Proof. Consider the problem LP_1 defined by (3.3) and (3.4). From the assumptions that $\mathrm{rank}(A) = m$ and problem (1.1) is feasible and nondegenerate, it is clear that problem LP_1 is nondegenerate. Therefore, if the interior point method is used to solve problem LP_1, then for $x > 0, t > 0$, which satisfy (3.4), $(AD_k^2A^T + t^2rr^T)^{-1}$ exists, where $r = b - Ax$.

It follows from (3.7) that $(AD_k^2A^T)^{-1}$ exists, too. As a result of the equivalence between algorithm A and the interior point method for solving problem LP_1, it is easy to see that for all $k, (AD_k^2A^T)^{-1}$ exists. Let E_k be the Euclidean norm about $x^{(k)}$. Then

$$(AD_k^2A^T)^{-1} = (A\tilde{D}_k^2E_k^2A^T)^{-1} = (A\tilde{D}_k^2A^T)^{-1}/E_k^2 \tag{3.13}$$

where $x^{(k)} = E_k\tilde{x}^{(K)}, \tilde{D}_k$ is a diagonal matrix with respect to $\tilde{x}^{(k)}$. Obviously, for all $k, (A\tilde{D}_k^2A^T)^{-1}$ exists, too. Thus, there must be a large M_1, such that for all k,

$$\|\tilde{D}_kA^T(A\tilde{D}_k^2A^T)^{-1}r\|_2 < M_1. \tag{3.14}$$

From (3.13) and (3.14), it is straightforward to establish

$$\begin{aligned}\|D_kA^T(AD_k^2A^T)^{-1}r\|_2 &= \|E_k\tilde{D}_kA^T(E_k^2A\tilde{D}_k^2A^T)^{-1}r\|_2 \\ &= \frac{1}{E_k}\|\tilde{D}_kA^T(A\tilde{D}_k^2A^T)^{-1}r\|_2 \le M_1/E\end{aligned}$$

where $E = \min_k\{E_k\}$. It follows from Lemma 3.4 that $E > 0$.

Let $M = M_1/E$. Then we have proved the lemma.

The proof of theorem 3.1 is as follows:

By (2.3) and (2.7),

$$\begin{aligned} r^{(k)} &= b - Ax^{(k)} = b - A[x^{(k-1)} + \alpha_{k-1}D_{k-1}p^{(k-1)}] \\ &= b - Ax^{(k-1)} - \alpha_{k-1}AD_{k-1}p^{(k-1)} = (1-\alpha_{k-1})r^{(k-1)} \\ &= \prod_{l=0}^{k-1}(1-\alpha_l)r^{(0)}. \end{aligned} \tag{3.15}$$

So substituting (3.15) into (2.4) gives

$$p^{(k)} = D_kA^T(AD_k^2A^T)^{-1}r^{(k)} = \prod_{l=0}^{k-1}(1-\alpha_l)D_kA^T(AD_k^2A^T)^{-1}r^{(0)}. \tag{3.16}$$

From Lemma 3.5, there exists a large $M > 0$, such that

$$\|D_kA^T(AD_k^2A^T)^{-1}r^{(0)}\|_2 < M. \tag{3.17}$$

Thus, by (3.16) and (3.17), there is an integer $k_0 > 0$, such that for all $k > k_0$,

$$\|p^{(k)}\|_2 \le M\prod_{l=0}(1-\alpha_l) < 1.$$

Therefore, as long as $k > k_0$, we have

$$-1 < P_k. \tag{3.18}$$

By (2.6) and (3.18), it is clear that

$$x^{(k+1)} = x^{(k)} + D_kp^{(k)} > 0,$$

and

$$Ax^{(k+1)} = Ax^{(k)} + AD_k p^{(k)} = Ax^{(k)} = b.$$

So, $x^{(k+1)}$ is a strictly feasible solution for problem (1.1).

This completes the proof of the theorem.

The proof of Theorem 3.2 is as follows:

First of all, we consider the sufficiency of the theorem. Assume that there exists an integer $k > 0$, such that (3.1) holds. Namely,

$$D_k A^T (AD_k^2 A^T)^{-1} r^{(k)} < 0, \quad b^T (AD_k^2 A^T)^{-1} . r^{(k)} > 0 \tag{3.19}$$

Let

$$y^{(k)} = (AD_k^2 A^T)^{-1} r^{(k)}.$$

Then, (3.19) can be rewritten as follows:

$$D_k A^T y^{(k)} \geq 0, \quad b^T y^{(k)} > 0.$$

Thus, it follows from the Farkas lemma that there is no feasible solution in problem (1.1).

The assertion of the necessity of the theorem is an immediate consequence of Lemma 3.3.

The theorem is proved.

Algorithm A described in Section 2 and the proof of its convergence can actually be considered a constructive proof for the Farkas lemma.

References

[1] M. S. Bazaraa and J. J. Jarvis, Linear Programming and Network Flows. John Wiley & Son, New York, 1977.

[2] P. E. Gill, W. Murray, M. A. Saunders, J. A. Tomlin and M. H. Wright, On Projected Newton Barrier Function Method for Linear Programming and an Equivalence to Karmarkar's Projected Method, SOL 85-11, Dept. of Operations Research, Stanford University, Stanford, CA 94305.

[3] M. Iri, and H. Imai, Multiplication Penalty Function Method for Linear Programming — Another New and Fast Algorithm, Dept. of Mathematical Enginerring and Instrumentation Physics, Faculty of Engineering, University of Tokyo, Japan, 1985.

[4] N. Karmarkar, A new polynomial-time algorithm for linear programming, combinatorica, 4 :4 (1984), 373–395.

[5] J. Ortega and W. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.

[6] R. J. Vanderbei, M. S. Meketon and B. A. Freadman, A Modification of Karmarkar's Linear Programming Algorithm, AT & T Bell-Lab, Holmdel, New Jersey, 1985.

[7] Wei Zi-luan, An exact solution to linear programming using an interior point method, *J. of Computational mathematics*, 5 : 3 (1987), 164–271.

[8] Wei Zi-luan, An interior point method for linear programming, *J. of Computational Mathematics*, 5 : 4 (1987), 342–351.