Journal of the Operations Research Society of Japan
Vol. 30, No. 2, June 1987

A PATH FOLLOWING ALGORITHM FOR STATIONARY POINT PROBLEMS

Yoshitsugu Yamamoto
University of Tsukuba

(Received April 18, 1986; Revised November 14, 1986)

Abstract We propose a path following algorithm for the stationary point problem: given a polytope $\Omega \subseteq R^n$ and an affine function $f: R^n \to R^n$ find a point $\hat{x} \in \Omega$ such that $\hat{x}\cdot f(\hat{x}) \leqq x\cdot f(\hat{x})$ for any point $x \in \Omega$. The linear system to be handled in the algorithm has only n+1 equations while the linear complementarity problem to which the problem is reduced has $n+m$ equations, where m is the number of constraints defining Ω. The algorithm is a variable dimension fixed point algorithm having as many rays as the vertices of Ω. It first leaves the starting point $w \in \Omega$ toward a vertex of Ω chosen by solving the linear programming problem: minimize $f(w)\cdot x$ subjects to $x \in \Omega$, and then moves on convex hulls of w and higher dimensional faces of Ω. Generally speaking, it terminates as soon as it hits the boundary of Ω or it finds a zero of f.

1. Introduction

Let $\Omega = \{ x : x \in R^n, a^i\cdot x \leqq b_i$ for $i \in M \}$ be a nonempty compact polyhedral set (a polytope) in R^n. For a given affine function $f(x) = Dx + c$ from R^n into R^n we consider the problem of finding a point $\hat{x} \in \Omega$ such that

(1.1) $\hat{x}\cdot f(\hat{x}) \leqq x\cdot f(\hat{x})$ for any $x \in \Omega$,

where $x\cdot f$ means inner product of x and f. The point $\hat{x} \in \Omega$ satisfying (1.1) is called a stationray point and the problem is called a stationary point problem. This problem arises from various fields such as quadratic programming, matrix game and economic equilibrium problem (see, for example Garcia and Zangwill [5]). It is known that the problem is cast into the linear complementarity problem, for which we have already several algorithms, e.g., Lemke [9], Reiser [11], Van der Heyden [6], Talman and Van der Heyden [12]. Eaves [3,4] adapted Lemke's algorithm for the stationary point problem and Pang [10] adapted the parametric principal pivoting algorithm. In van der Laan and Talman [8] a variable dimension algorithm was adapted for the linear complementarity problem with upper and lower bounds, which can be considered to be a stationray point problem with $\Omega = \{ x : x \in R^n, a \leqq x \leqq b \}$. Eaves' algorithm indeed deals with a general convex polyhedral set Ω but it must handle a system of $n+m$ linear equations, where $m = |M|$. On the other hand, the linear system in van der Laan and Talman's algorithm has only n equations. But it makes use of the trivial face structure of Ω. Recently Talman and Yamamoto [13] developed an algorithm for stationary point problems with a nonlinear function f. Assuming that Ω is a simple polytope and its face structure is known in advance, they made a triangulation of Ω and proposed a simplicial algorithm based on the piecewise linear approximation of f.

In this paper we propose an algorithm for stationary point problems which does not assume the knowledge of the face structure of Ω and handles a system of $n+1$ linear equations. When the starting point w is in the relative interior of Ω, the algorithm can be viewed as a variable dimension algorithm with as many rays as vertices of Ω. At the start it solves the linear programming problem

$$\text{minimize} \quad f(w)\cdot x \quad \text{subject to} \quad x \in \Omega$$

to obtain an optimum vertex v and leaves point w toward v with hope that the function value does not change quickly and so vertex v remains an optimum solution of

$$\text{minimize} \quad f(v)\cdot x \quad \text{subject to} \quad x \in \Omega.$$

If it is the case, vertex v is a stationary point. If not, the algorithm changes the direction according to the function value at some point between w and v. It terminates as soon as either it hits the boundary of Ω or finds a zero of f, a trivial stationray point, in Ω.

The organization of this paper is as follows. In Section 2 we review subdivided manifolds, a basic theorem for fixed point algorithms and the primal-dual pair of subdivided manifolds. Based on these preliminaries we prove in Section 3 that there is a finite path of solutions of a certain system of equations connectiong the starting point and a stationary point. In Section 4 we describe the algorithm and show that it traces the path and then terminates after a finite number of iterations with a stationray point. Some remarks are found in Section 5.

2. Basic Theorem for Fixed Point Algorithms and Primal-Dual Pair of Subdivided Manifolds

We give a brief review of a basic theorem for fixed point algorithms and the primal-dual pair of subdivided manifolds introduced by Kojima and Yamamoto [7] as a unifying framework for a class of fixed point algorithms.

We call a convex polyhedral set a cell or an ℓ-cell to clarify it dimension. When a cell B is a face of a cell C, we write $B < C$. Especially when B is a facet of C, we write $B \triangleleft C$.

Let $\mathcal{M}$ be a finite or countable collection of ℓ-cells. We write $\overline{\mathcal{M}} = \{ B : B < C \text{ for some } C \in \mathcal{M} \}$ and $|\mathcal{M}| = \bigcup \{ C : C \in \mathcal{M} \}$. We call $\mathcal{M}$ a subdivided ℓ-manifold if and only if

(2.1) for any $B, C \in \mathcal{M}$, $B \cap C = \emptyset$ or $B \cap C < B$ and C,

(2.2) for each $(\ell-1)$-cell B of $\overline{\mathcal{M}}$ at most two ℓ-cells of $\mathcal{M}$ have B as a facet,

(2.3) $\mathcal{M}$ is locally finite : each point $x \in |\mathcal{M}|$ has a neighborhood which intersects only a finite number of cells of $\mathcal{M}$.

We write $\partial\mathcal{M} = \{ B : B \in \overline{\mathcal{M}}, B \triangleleft C \text{ for exactly one } \ell\text{-cell } C \text{ of } \mathcal{M} \}$ and call it the boundary of $\mathcal{M}$.

A continuous function $h : |\mathcal{M}| \to R^k$ is said to be a piecewise linear (pl for short) function on $\mathcal{M}$ if the restriction of h to each cell of $\mathcal{M}$ is an affine function. For a subdivided $(n+1)$-manifold $\mathcal{M}$ and a pl function $h : |\mathcal{M}| \to R^n$ we say that $r \in R^n$ is a regular value of h if $B \in \overline{\mathcal{M}}$ and $h^{-1}(r) \cap B \neq \emptyset$ imply that $\dim h(B) = n$. The following theorem is a basic theorem for fixed point algorithms (see Eaves [2]).

Theorem 2.1. Let $\mathcal{M}$ be a subdivided $(n+1)$-manifold, $h : |\mathcal{M}| \to R^n$ be a pl function. Suppose $r \in R^n$ is a regular value of h. Then $h^{-1}(r)$ is a disjoint union of paths and loops, where a path is a subdivided 1-manifold homeomorphic to one of the intervals $(0,1)$, $(0,1]$ and $[0,1]$ and a loop is a subdivided 1-manifold homeomorphic to the 1-dimensional sphere. Furthermore they satisfy the following conditions.

(2.4) $h^{-1}(r) \cap C$ is either empty or a 1-cell for each $C \in \mathcal{M}$.

(2.5) A loop of $h^{-1}(r)$ does not intersect $|\partial\mathcal{M}|$.

(2.6) If a path S of $h^{-1}(r)$ is compact, ∂S consists of two distinct points in $|\partial\mathcal{M}|$.

Let $\mathcal{P}$ and $\mathcal{D}$ be subdivided manifolds. If $\mathcal{P}$ and $\mathcal{D}$ satisfy the following conditions with some positive integer ℓ and an operator $d : \overline{\mathcal{P}} \cup \overline{\mathcal{D}} \to \overline{\mathcal{P}} \cup \overline{\mathcal{D}} \cup \{\emptyset\}$, we say that $(\mathcal{P},\mathcal{D};d)$ is a primal-dual pair of subdivided manifolds (PDM for short) with degree ℓ.

(2.7) For each $X \in \overline{\mathcal{P}}$ $X^d \in \overline{\mathcal{D}} \cup \{\emptyset\}$ and for each $Y \in \overline{\mathcal{D}}$ $Y^d \in \overline{\mathcal{P}} \cup \{\emptyset\}$.

(2.8) If $Z \in \overline{\mathcal{P}} \cup \overline{\mathcal{D}}$ and $Z^d \neq \emptyset$, then $(Z^d)^d = Z$ and $\dim Z + \dim Z^d = \ell$.

(2.9) If $Z_1, Z_2 \in \overline{\mathcal{P}}$ (or $\overline{\mathcal{D}}$), $Z_1 < Z_2$, $Z_1^d \neq \emptyset$ and $Z_2^d \neq \emptyset$, then $Z_2^d < Z_1^d$.

We call the operator d the dual operator and Z^d the dual of Z.

For a PDM $(\mathcal{P},\mathcal{D};d)$ with degree ℓ let

$$\langle\mathcal{P},\mathcal{D};d\rangle = \{ X \times X^d : X \in \overline{\mathcal{P}},\ X^d \neq \emptyset \},$$

or equivalently

$$\langle\mathcal{P},\mathcal{D};d\rangle = \{ Y^d \times Y : Y \in \overline{\mathcal{D}},\ Y^d \neq \emptyset \}.$$

Then we have the following theorem. See Kojima and Yamamoto [7] for the proof.

Theorem 2.2. Let $(\mathcal{P},\mathcal{D};d)$ be a PDM with degree ℓ. Then $L = \langle\mathcal{P},\mathcal{D};d\rangle$ is a subdivided ℓ-manifold and

$$\partial L = \{ X \times Y : X \times Y \text{ is an } (\ell-1)\text{-cell of } \overline{L},\ X \in \overline{\mathcal{P}},\ Y \in \overline{\mathcal{D}}, \text{ and either } X^d = \emptyset \text{ or } Y^d = \emptyset \}.$$

3. Basic Model of the Algorithm

Let $\mathcal{F}$ be the family of all faces of Ω. For each face $F \in \mathcal{F}$ let

$$I(F) = \{ i : i \in M,\ a^i \cdot x = b_i \text{ for any point } x \in F \},$$

the index set of active constraints at face F, and F^* be the cone generated by a^i's for $i \in I(F)$, i.e.,

$$F^* = \{ y : y = \sum_{i \in I(F)} \mu_i a^i,\ \mu_i \geq 0 \text{ for any } i \in I(F) \},$$

where we assume that $F^* = \{0\}$ when $I(F) = \emptyset$. Cone F^* is called the dual cone of face F. Note that $\dim F^* = n - \dim F$ and Ω^* is the orthogonal complement of the tangential space of Ω. Then the stationary point problem is a problem of finding a point $x \in \Omega$ and a face $F \in \mathcal{F}$ such that

(3.1) $x \in F$ and $-f(x) \in F^*$.

We show three different stationary points in Fig. 1, where F_1 is a zero dimensional face consisting of point x^1 and F_3 is Ω itself. Point x^3 is also a stationary point because $f(x^3) = 0$.

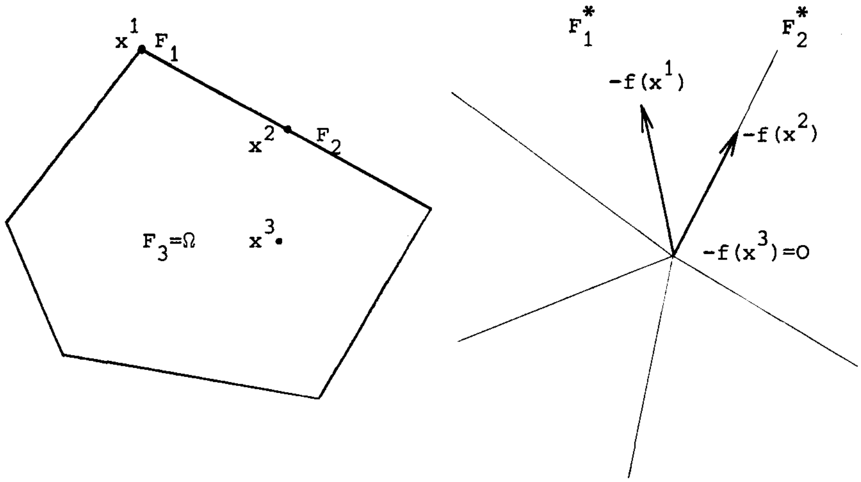

Fig. 1. Stationary points, faces and dual cones

The key point for developing a path following algorithm for the stationary point problem is to construct a subdivided manifold $\mathcal{L}$ such that $\partial\mathcal{L}$ has a trivial starting point and $F \times F^*$ for all faces F of Ω.

Let $w \in \Omega$ be an initial guess of a stationary point. We do not require point w to lie in the relative interior of Ω. For each $F \in \mathcal{F}$ with $w \notin F$ let wF be the join of point w and face F, i.e.,

$$wF = \{ x : x = \alpha w + (1-\alpha)z \text{ for some } z \in F \text{ and } 0 \leqq \alpha \leqq 1 \}.$$

Note that $\dim wF = \dim F + 1$. Let

(3.2) $\mathcal{P} = \{ wF : w \notin F \in \mathcal{F}, \dim F = \dim \Omega - 1 \}.$

Examples of $\mathcal{P}$ are shown in Fig. 2 for some different starting points. Then $\mathcal{P}$ is a subdivided manifold of the same dimension as Ω and

(3.3.a) $\overline{\mathcal{P}} = \{ wF : w \notin F \in \mathcal{F} \} \cup \{ F : w \notin F \in \mathcal{F} \} \cup \{\{w\}\},$

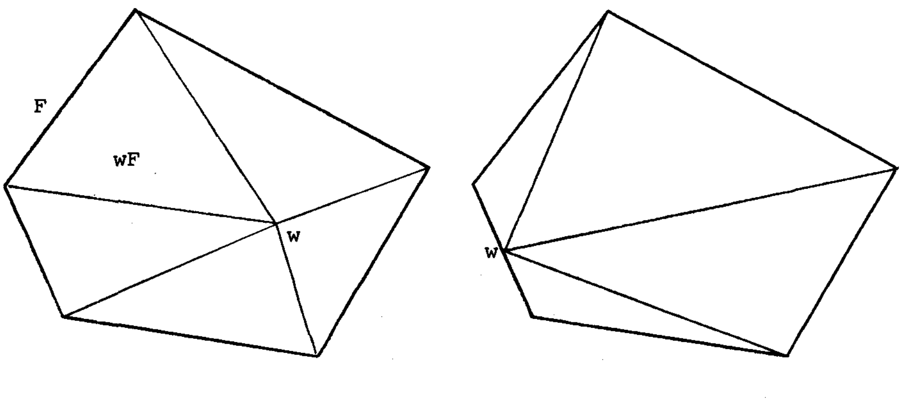

Fig. 2. Primal subdivided manifold $\mathcal{P}$

(3.3.b) $|\mathcal{P}| = \Omega$.

Let $\mathcal{D}$ be the collection of dual cones of all vertices of Ω (see Fig.3), i.e.,

(3.4) $\mathcal{D} = \{ F^* : F \in \mathcal{F},\ \dim F = 0 \}$.

Then $\mathcal{D}$ is a subdivided n-manifold and

(3.5) $\bar{\mathcal{D}} = \{ F^* : F \in \mathcal{F} \}$,
$|\mathcal{D}| = R^n$.

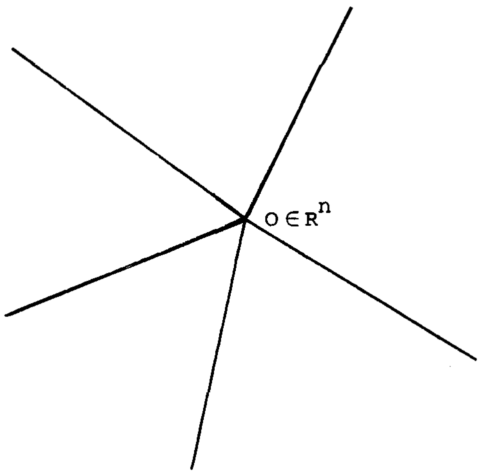

Fig. 3. Dual subdivided manifold $\mathcal{D}$

Now let the dual operator d be

$$
\begin{aligned}
(wF)^d &= F^* &&\text{if } w \notin F \in \mathcal{F} \\
F^d &= \emptyset &&\text{if } w \notin F \in \mathcal{F} \\
\{w\}^d &= \emptyset && \\
(F^*)^d &= wF &&\text{if } w \notin F \in \mathcal{F} \\
&= \emptyset &&\text{if } w \in F \in \mathcal{F}.
\end{aligned} \tag{3.6}
$$

We readily see that $(\mathcal{P},\mathcal{D};d)$ is a PDM with degree $n+1$ and we obtain the following lemma from Theorem 2.2.

Lemma 3.1. Let $(\mathcal{P},\mathcal{D};d)$ be a PDM with degree $n+1$ defined by (3.2), (3.4) and (3.6) and let $\mathcal{L} = \langle \mathcal{P},\mathcal{D};d \rangle$. Then

(3.7) $\mathcal{L}$ is a subdivided (n+1)-manifold.

(3.8) $\partial\mathcal{L} = \{\{w\} \times F^* : w \notin F \in \mathcal{F},\ \dim F = 0\} \cup \{ F \times F^* : w \notin F \in \mathcal{F} \}$
$\cup \{ wE \times F^* : w \in F \in \mathcal{F},\ w \notin E \lhd F,\ \dim F > 0 \}$.

(3.9) $|\partial L| = (\{w\} \times \bigcup\{ F^* : w \notin F \in F,\ \dim F = 0 \})$
$\cup\ (\bigcup\{ F \times F^* : F \in F \})$.

proof. (3.7) is a direct consequence of Theorem 2.2. We prove (3.8). Suppose $X \times Y \in \partial L$. Then $\dim X + \dim Y = n$ and either X^d or Y^d is empty. If $X^d = \emptyset$ (resp. $Y^d = \emptyset$), then the unique cell of L having $X \times Y$ is $Y^d \times Y$ (resp. $X \times X^d$). Suppose first $X^d = \emptyset$. Then $X = \{w\}$ or F such that $w \notin F \in F$. When $X = \{w\}$, we have $\dim Y = n$ and $\{w\} \lhd Y^d$. This implies that $Y = F^*$ and $\dim F = n - \dim F^* = 0$. When $X = F$ with $w \notin F \in F$, we have $\dim Y = n - \dim F$, $F \lhd Y^d$. Therefore $Y = F^*$. Next suppose $Y^d = \emptyset$, i.e., $Y = F^*$ for some $w \in F \in F$. Then $\dim X = n - \dim F^* = \dim F$ and $F^* \lhd X^d$. Therefore we obtain that $X = wE$ for some $E \in F$ such that $\dim E = \dim F - 1$, $w \notin E$ and $F^* \lhd E^*$. By the inclusion reversing property of F and F^*, we see $E \lhd F$.

Since all the cells above clearly belong to ∂L, we have proved (3.8).

By noting that $\bigcup\{ wE : w \notin E \lhd F \} = F$ we obtain (3.9) from (3.8). //

Now let a pl function $h : |L| \to R^n$ be defined by

$$h(x,y) = y + f(x) \quad \text{for} \quad (x,y) \in |L|$$

and consider the system of equations

(3.10) $h(x,y) = 0, \quad (x,y) \in |L|$.

Note that the system has n equations and $2n$ variables, which are, however, restricted to $(n+1)$-dimensional subdivided manifold L. By applying Theorem 2.1 to (3.10) we have the following theorem.

Theorem 3.2. Suppose that $w \in \Omega$ is not a stationary point and $0 \in R^n$ is a regular value of $h : |L| \to R^n$. Then

(3.11) $(x^o,y^o) = (w,-f(w))$ lies in $h^{-1}(0) \cap |\partial L|$, and

(3.12) there is a path S of $h^{-1}(0)$ from (x^o,y^o) to a point $(x,y) \in |\partial L|$ such that x is a stationary point and $y = -f(x)$. Furthermore S consists of a finite number of line segments.

proof. Since w is not a stationary point, $-f(w)$ does not lie in F^* for any face $F \in F$ having w. Therefore by (3.5) $-f(w) \in F^*$ for some $F \in F$ such that $w \notin F$ and $\dim F = 0$. This and (3.9) prove (3.11).

If $0 \in R^n$ is a regular value of $h : |L| \to R^n$, we can apply Theorem 2.1 to h. By (2.5) the connected component S of $h^{-1}(0)$ having (x^o, y^o) is a path. Suppose S intersects a cell $C = wF \times F^* \in L$. Since wF is compact and $S \cap C \subseteq wF \times \{ -f(x) : x \in wF \}$, $S \cap C$ is a compact line segment. By the definition of L, it consists of finite cells. Therefore S consists of finite line segments and hence S is compact. Then by (2.6) S has another end-point, say (x,y), in $|\partial L|$. Suppose first $(x,y) \in \{w\} \times \bigcup\{ F^* : w \notin F \in F, \dim F = 0 \}$. Then $x = w$. Since $(x,y) \in h^{-1}(0)$, $y = -f(x) = -f(w) = y^0$. Hence we have $(x,y) = (x^0,y^0)$, which contradicts (2.6). Therefore by (3.9) we have $(x,y) = (x,-f(x)) \in F \times F^*$ for some $F \in F$. By (3.1) we have that x is a stationary point. //

Theorem 3.2 shows that we will find a stationary point x by tracing the finite path S from the trivial starting point $(x^0,y^0) = (w,-f(w))$. In the next section we show that given an end-point of a line segment of S the optimum solution of a linear programming problem gives the other end-point, which serves as an initial end-point of the next line segment.

4. The Algorithm

Now suppose that $h^{-1}(0)$ intersects a cell $C = wF \times F^* \in L$. Let $U(F)$ be the set of all vertices of face F. Then wF is the convex hull of w and $U(F)$. Therefore $h^{-1}(0) \cap C \neq \emptyset$ if and only if the following system (4.1) has a solution (λ,μ).

$$
(4.1)\qquad
\begin{aligned}
&\sum_{i\in I(F)} \mu_i a^i + \sum_{u\in U(F)} \lambda_u Du + \lambda_w Dw = -c \\
&\sum_{u\in U(F)} \lambda_u + \lambda_w = 1 \\
&\lambda_u \geq 0 \quad \text{for all} \quad u \in U(F), \quad \lambda_w \geq 0 \\
&\mu_i \geq 0 \quad \text{for all} \quad i \in I(F).
\end{aligned}
$$

Clearly $(x,y) \in h^{-1}(0) \cap C$ is given by

$$
(4.2)\qquad x = \sum_{u\in U(F)} \lambda_u u + \lambda_w w, \qquad y = \sum_{i\in I(F)} \mu_i a^i.
$$

Note that the set of solutions of (4.1) is generally unbounded. For example, if $a^i = -a^j$ for some $i, j \in I(F)$, it is clearly unbounded. However, as shown in the proof of Theorem 3.2, the set $h^{-1}(0) \cap C$ is a line segment and we have the following lemma.

Lemma 4.1. Suppose $0 \in R^n$ is a regular value of $h : |L| \to R^n$ and $h^{-1}(0) \cap C \neq \emptyset$ for cell $C \in L$. Then for any vector $(p,q) \in R^{2n}$ the linear programming problem of minimizing or maximizing

$$
(4.3)\qquad p\cdot x + q\cdot y \quad \text{subject to (4.1) and (4.2)}
$$

has an optimum solution. Let $(\lambda^1,\mu^1,x^1,y^1)$ and $(\lambda^2,\mu^2,x^2,y^2)$ be a minimizer and a maximizer of the linear programming problem. If $p\cdot x^1 + q\cdot y^1 \neq p\cdot x^2 + q\cdot y^2$, then (x^1,y^1) and (x^2,y^2) are two distinct end-points of line segement $h^{-1}(0) \cap C$.

Note that we do not have to know all the vertices of $U(F)$ in advance when we solve the linear programming problem (4.3). The set $U(F)$ is the set of vertices of the face $F = \{ x : x \in \Omega,\ a^i\cdot x = b_i \text{ for } i \in I(F) \}$ and hence necessary vertices of $U(F)$ can be generated at need. See, for example, Chapter 23 on the deconposition principle in Dantzig [1]. Therefore we have only to know the index set $I(F)$ and to make an appropriate objective function to guarantee the condition in Lemma 4.1. The next lemma gives us how to find $I(F)$ for face F. Here we abbreviate $I(\{u\})$ by $I(u)$ when u is a point.

Lemma 4.2. Let $u^1,\ldots, u^k$ be points on face F such that the affine hull $\mathrm{aff}(\{ u^j : j=1,\ldots,k \})$ of the points contains F. Then

$$I(F) = \bigcap_{j=1}^{k} I(u^j).$$

proof. Let i be an arbitrary index of $\bigcap_{j=1}^{k} I(u^j)$. Since any point x of F is an affine combination of $u^1,\ldots,u^k$,

$$a^i\cdot x = a^i\cdot(\sum_{j=1}^{k} \alpha_j u^j) = \sum_{j=1}^{k} \alpha_j b_i = b_i.$$

This means that $i \in I(F)$. We have seen that $\bigcap_{j=1}^{k} I(u^j) \subseteq I(F)$. Since the reverse relation is trivial, we have proved the lemma. //

For a subset J of M we write

$$F(J) = \{ x : x \in \Omega,\ a^i\cdot x = b_i \text{ for } i \in J \}.$$

Note that several different subsets of M may define the same face F and further they may be different from $I(F)$. Index set $I(F)$ is the maximum subset of M defining face F. In fact $I(F(J))$, the index set of active constraints of a face defined by $J \subseteq M$, may not coincide with J.

The following lemma gives us the termination condition of the algorithm.

Lemma 4.3. Let (λ,μ) be a solution of (4.1) and let

$$I_+ = \{ i : i \in I(F) \text{ and } \mu_i > 0 \}.$$

If either

(4.4) $\lambda_w = 0$ or $I_+ \subseteq I(w)$,

then $x = \sum_{u\in U(F)} \lambda_u u + \lambda_w w$ is a stationary point.

proof. Since $I_+ \subseteq I(F)$, $F \subseteq F(I_+)$ in general. If $\lambda_w = 0$, point $x = \Sigma_{u \in U(F)} \lambda_u u \in F \subseteq F(I_+)$. If $I_+ \subseteq I(w)$, then $w \in F(I_+)$ and hence we have $x \in F(I_+)$. Therefore in either case $x \in F(I_+)$ and $-f(x) = \Sigma_{i \in I_+} \mu_i a^i \in F(I_+)^*$. //

Now we give the algorithm. Here $A(I,U)$ denotes the set of (λ,μ) satisfying (4.1) with $I(F)$ and $U(F)$ replaced by I and U, respectively. We also abbreviate $U(F(I))$ by $U(I)$ and the linear subspace spanned by a set A of vectors by $spc(A)$.

```
     program stationary;
     begin
     read (w);
{1}  v:=argmin{f(w)·x : x ∈ Ω};
{2}  if v=w then
       writeln('starting point w is a stationary point.')
     else
       begin
{3}    I:=I(v); U:={v}; f:=0;
{4}    (λ,μ):=argmin{-λ_v : (λ,μ) ∈ A(I,U)};
       I_+:={i : i ∈ I, μ_i > 0}; U_+:={u : u ∈ U,λ_u > 0};
{5}    while λ_w≠0 and I_+ ⊈ I(w) do
         begin
{6}      if |U_+|=f+1 then
           begin
{7}        choose k ∈ I \ I_+ such that a^k is linearly independent of a^i
           (i ∈ I_+);
{8}        find v ∈ U(I_+) such that a^k·v < b_k;
{9}        I:=I ∩ I(v); U:=U(I); f:=f+1;
{10}       find p ∈ aff(U_+ ∪ {v} ∪ {w}) such that p·(u-w)=0 for u ∈ U_+ and
           p·(v-w) < 0;
{11}       (λ,μ,x):=argmin{p·x : x=Σ_{u∈U}λ_u u+λ_w w, (λ,μ) ∈ A(I,U)};

           I_+:={i : i ∈ I, μ_i > 0}; U_+:={u : u ∈ U, λ_u > 0};
           end
         else
           begin
{12}       I':= ∩_{u∈U_+} I(u);
{13}       choose k ∈ I' \ I;
{14}       I:=I'; U:=U(I); f:=f-1;
{15}       find q ∈ spc({a^i : i ∈ I_+} ∪ {a^k}) such that q·a^i=0 for i ∈ I_+
           and q·a^k < 0;
{16}       (λ,μ,y):=argmin{q·y : y=Σ_{i∈I}μ_i a^i, (λ,μ) ∈ A(I,U)};
           I_+:={i : i ∈ I, μ_i > 0}; U_+:={u : u ∈ U, λ_u > 0}
           end
         end
```

```
{17}    x:=Σ_{u∈U} λ_u u+λ_w w;
        writeln('A stationary point is found.',x)
        end
      end.
```

We give several lemmas to show that the algorithm traces tha path S. For the time being we assume that the system (4.1) is nondegenerate. Note that the nondegeneracy assumption implies the regular value assumption of $0 \in R^n$ for $h : |L| \to R^n$ but the converse is not true. If there is an $i \in I$ such that a^i is a nonnegative combination of some other a^j's with $j \in I$, the system (4.1) may be degenerate although $0 \in R^n$ is a regular value of h. This is due to the fact that the regular value assumption is based on the face structure of $F(I)^*$ and does not care whether a higher dimensional face of $F(I)^*$ has a vector a^i in its interior.

Through the following lemmas we assume that

$$\lambda_w > 0$$

and employ the notations

$$(4.5)\qquad \begin{aligned} x &= \sum_{u\in U(J)} \lambda_u u + \lambda_w w \\ y &= \sum_{i\in J} \mu_i a^i \\ U_+ &= \{ u : u \in U(J),\ \lambda_u > 0 \} \\ J_+ &= \{ i : i \in J,\ \mu_i > 0 \}. \end{aligned}$$

Lemma 4.4. Let (λ,μ) be a point of $A(J,U(J))$. Suppose

(a) (x,y) defined by (4.5) is an end-point of line segment $S \cap (wF(J) \times F(J)^*)$ and lies in facet $wF(J) \times F(I)^*$ of both cells $wF(J) \times F(J)^*$ and $wF(I) \times F(I)^*$,

(b) $F(J) \subseteq \mathrm{aff}(U_+)$,

(c) $p \in R^n$ satisfies the condition in Step 10 for some vertex $v \in U(I) \setminus U(J)$.

Let (λ',μ',x',y') be a minimizer of $p\cdot x$ under $(\lambda,\mu) \in A(I,U(I))$ and (4.5). Then (x',y') is the opposite end-point of line segment $S \cap (wF(I) \times F(I)^*)$.

proof. Since $x \in wF(J) \subseteq \mathrm{aff}(U_+ \cup \{w\})$, $p\cdot x = p\cdot w$ by the choice of p. Since $U(I)$ has a vertex v satisfying $p\cdot v < p\cdot w$, $wF(I)$ is not contained in the plane $\{ x : p\cdot x = p\cdot w \}$ but lies in the half space $\{ x : p\cdot x \leqq p\cdot w \}$. Therefore (λ,μ,x,y) is a maximizer of $p\cdot x$ under

$(\lambda,\mu) \in A(I,U(I))$ and (4.5). Since the system (4.1) is nondegenerate and $v \in U(I)$, it is clear that $p\cdot x' < p\cdot w$. Therefore by Lemma 4.1 we have the desired result. //

By almost the same argument we have the following lemma.

Lemma 4.5. Let (λ,μ) be a point of $A(J,U(J))$. Suppose
(a) (x,y) defined by (4.5) is an end-point of line segment $S \cap (wF(J) \times F(J)^*)$ and lies in facet $wF(I) \times F(J)^*$ of both cells $wF(J) \times F(J)^*$ and $wF(I) \times F(I)^*$,
(b) $F(J)^* \subseteq \mathrm{spc}(\{a^i : i \in J_+\})$,
(c) $q \in R^n$ satisfies the condition in Step 15 for some $k \in I \setminus J$.
Let (λ',μ',x',y') be a minimizer of $q\cdot y$ under $(\lambda,\mu) \in A(I,U(I))$ and (4.5). Then (x',y') is the opposite end-point of line segement $S \cap (wF(I) \times F(I)^*)$.

The following lemma gives us the facet which has a given end-point of line segment $S \cap (wF(J) \times F(J)^*)$.

Lemma 4.6. Let (λ,μ) be a basic solution of $A(J,U(J))$ such that (x,y) defined by (4.5) is an end-point of line segment $S \cap (wF(J) \times F(J)^*)$. If $|U_+| = \dim F(J) + 1$, then point (x,y) lies in $wF(J) \times F(I)^*$ for some $F(I)^* \triangleleft F(J)^*$. If $|U_+| < \dim F(J) + 1$, then (x,y) lies in $wF(I) \times F(J)^*$ for some $F(I) \triangleleft F(J)$.

proof. Since (x,y) is an end-point of $S \cap (wF(J) \times F(J)^*)$, it lies on the boundary of $wF(J) \times F(J)^*$. Therefore it is either in $wF(J) \times F(I)^*$ for some $F(I)^* \triangleleft F(J)^*$ or in $wF(I) \times F(J)^*$ for some $F(I) \triangleleft F(J)$. Note that w and U_+ are affinely independent because λ_w and λ_u are basic variables for $u \in U_+$. Therefore if $|U_+| = \dim F(J) + 1$, point $x = \Sigma_{u\in U_+}\lambda_u u + \lambda_w w$ is in the relative interior of $wF(J)$. This implies that $(x,y) \in wF(J) \times F(I)^*$. If $|U_+| < \dim F(J) + 1$, $|J_+| = (n+1) - (|U_+|+1) = n - |U_+| \geqq n - \dim F(J) = \dim F(J)^*$ because $\lambda_w > 0$ and the system (4.1) is nondegenerate. Note that a^i's $(i \in J_+)$ are linearly independent because μ_i is a basic variable for $i \in J_+$. Therefore $y = \Sigma_{i\in J_+}\mu_i a^i$ is in the relative interior of $F(J)^*$. Thus $(x,y) \in wF(I) \times F(J)^*$. //

Lemma 4.7. Let (λ,μ) be a basic solution of $A(J,U(J))$ such that (x,y) defined by (4.5) is an end-point of line segment $S \cap (wF(J) \times F(J)^*)$. Suppose

(a) $J = I(F(J))$,

i.e., set J is the maximum index set defining $F(J)$, and

(b) $|U_+| = \dim F(J) + 1$.

Then

(4.6) there exists a $k \in J \setminus J_+$ such that a^i's ($i \in J_+$) and a^k are linearly independent,

(4.7) there is a vertex $v \in U(J_+)$ such that $a^k \cdot v < b_k$, and

(4.8) let $I = J \cap I(v)$, then $I = I(F(I))$, $F(J) \triangleleft F(I)$ and $\dim F(I) = \dim F(J) + 1$.

proof. In general $|J| \geq n - \dim F(J)$. By the assumptions $\lambda_w > 0$, (b) and nondegeneracy, we have $|J_+| = (n+1) - (|U_+|+1) < n - \dim F(J)$. Therefore $J \setminus J_+ \neq \emptyset$. Suppose a^k is linearly dependent on a^i's ($i \in J_+$) for any $k \in J \setminus J_+$. Then

$$n - \dim F(J) = \dim F(J)^* = \text{rank } \{ a^i : i \in J \}$$
$$= \text{rank } \{ a^i : i \in J_+ \} \leq |J_+| < n - \dim F(J).$$

This is a contradiction and we have (4.6).

To see (4.7) note first that $U(J_+) \neq \emptyset$ because $F(J_+)$ is nonemtpy and bounded. Suppose that $a^k \cdot u = b_k$ for any $u \in U(J_+)$. Since a^i's ($i \in J_+$) and a^k are linearly independent by (4.6), $\dim F(J_+) \leq n - (|J_+|+1) < n - |J_+|$. This is a contradiction and we have (4.7).

Since $a^k \cdot u = b_k$ for any $u \in U_+$, U_+ and vertex v in (4.7) are affinely independent. We first show that $F(J_+) \subseteq \text{aff}(U_+ \cup \{v\})$. Since $v \in U(J_+) \subseteq F(J_+)$ and $U_+ \subseteq U(J) \subseteq F(J) \subseteq F(J_+)$, we have $U_+ \cup \{v\} \subseteq F(J_+)$. Furthermore, since a^i's ($i \in J_+$) are linearly independent, $\dim F(J_+) = n - |J_+| = |U_+|$. Therefore we have $F(J_+) \subseteq \text{aff}(U_+ \cup \{v\})$. And by Lemma 4.2 $I(F(J_+)) = (\bigcap_{u \in U_+} I(u)) \cap I(v)$. By (a), (b) and Lemma 4.2 we also have $J = \bigcap_{u \in U_+} I(u)$. Therefore again by Lemma 4.2 $I(F(J_+)) = J \cap I(v) = I(F(I))$. Since $J_+ \subseteq J$ and $\dim F(J_+) = |U_+| = \dim F(J) + 1$, we have $F(J) \lessdot F(J_+) = F(I)$. //

Lemma 4.8. Let (λ, μ) be a basic solution of $A(J, U(J))$ such that (x,y) defined by (4.5) is an end-point of line segment $S \cap (wF(J) \times F(J)^*)$. Suppose

(a) $J = I(F(J))$

and

(b) $|U_+| < \dim F(J) + 1$.

Let $I' = \bigcap_{u \in U_+} I(u)$, then

(4.9) $I' \setminus J \neq \emptyset$,

(4.10) a^i's ($i \in J_+$) and a^k are linearly independent for any $k \in I'$

$\setminus J$,

(4.11) $I' = I(F(I'))$, $F(I') \lhd F(J)$ and $\dim F(I') = \dim F(J) - 1$.

proof. By the assumptions $\lambda_w > 0$, (b) and that (x,y) is an end-point of the line segment, we see that U_+ is contained in some facet, say F', of $F(J)$. Therefore $\emptyset \neq I(F') \setminus J \subseteq \bigcap_{u \in U_+} I(u) \setminus J$, which implies (4.9).

To show (4.10) suppose the contrary, i.e., $a^k = \Sigma_{i \in J_+} \alpha_i a^i$ for some α_i ($i \in J_+$). Since for any vertex u of U_+ $a^i \cdot u = b_i$ for $i \in J_+$ and $a^k \cdot u = b_k$, $b_k = \Sigma_{i \in J_+} \alpha_i b_i$. Therefore for any vertex u of $U(J)$, which is a subset of $U(J_+)$, $a^k \cdot u = b_k$. This and (a) imply that $k \in J$, which is a contradiction. Hence we have (4.10).

Since a^i's ($i \in J_+$) and a^k are linearly independent, $\dim F(J_+ \cup \{k\}) = n - (|J_+|+1) = |U_+| - 1$. Note that $U_+ \subseteq F(J \cup \{k\}) \subseteq F(J_+ \cup \{k\})$, then we have $F(J_+ \cup \{k\}) \subseteq \operatorname{aff} U_+$. Therefore by Lemma 4.2 $F(J_+ \cup \{k\}) = F(\bigcap_{u \in U_+} I(u)) = F(I')$ and $I(F(I')) = I'$.

In general $\dim F(J) = n - \dim F(J)^* \leqq n - |J_+|$. Suppose that $\dim F(J) \leqq n - |J_+| - 1$, then by (b) $\dim F(J) \leqq |U_+| - 1 < \dim F(J)$, a contradiction. Therefore we have $\dim F(J) = n - |J_+| = |U_+|$. since we have seen that $\dim F(I') = \dim F(J_+ \cup \{k\}) = |U_+| - 1$, we obtain $F(I') \lhd F(J)$. //

The following two lemmas guarantee Steps 10 and 15.

Lemma 4.9. Let $U_+ = \{u^1, \ldots, u^\ell\}$. If U_+, v and w are affinely independent, the vector $p \in R^n$ in Step 10 is obtained by $p = P(P^tP)^{-1}e$, where $P = [u^1-w, \ldots, u^\ell-w, v-w]$ and $e = (0,\ldots,0,-1)^t \in R^{\ell+1}$.

Lemma 4.10. Let $I_+ = \{i_1, \ldots, i_\ell\}$. If a^i's ($i \in I_+$) and a^k are linearly independent, the vector $q \in R^n$ in Step 15 is obtained by $q = Q(Q^tQ)^{-1}e$, where $Q = [a^{i_1}, \ldots, a^{i_\ell}, a^k]$ and $e = (0,\ldots,0,-1)^t \in R^{\ell+1}$.

Combining above lemmas we obtain the following theorem.

Theorem 4.11. The algorithm finds a stationary point after finitely many iterations.

proof. If $v = w$ in Step 2, starting point w is a stationary point. Otherwise, $(x^0,y^0) = (w,-f(w))$ lies in $S \cap |\partial L|$ by (3.11). Note that $I = I(F(I))$ and $f = \dim F(I)$ in Step 3 and the algorithm traces path S in $wF(I) \times F(I)^* = w\{v\} \times \{v\}^*$ by minimizing $-\lambda_v$ in Step 4. Then the theorem follows Lemma 4.2 to 4.10 by induction. //

5. Remarks

We give several remarks in order.

(i) If we solve the dual problem

$$\text{minimize} \sum_{i \in M} b_i \mu_i \quad \text{subject to} \quad (\lambda, \mu) \in A(M, \emptyset),$$

of the linear programming problem to be solved in Step 1, we obtain the initial basic solution of system (4.1).

(ii) The column vector a^k independent of a^i's ($i \in I_+$) is easily found in Step 7 if we keep the inverse of the basic matrix of (4.1).

(iii) Through the algorithm we solve a series of linear programming problems, each of which differs slightly from the former. It should be noted that we always have a feasible basic solution at hand for the new linear programming problem.

(iv) The column generation technique is used in Steps 11 and 16. Therefore we have a feasible basic solution of the system

$$\begin{aligned} a^i \cdot x + s_i &= b_i \quad \text{for} \quad i \in M \setminus I \\ a^i \cdot x \quad &= b_i \quad \text{for} \quad i \in I \\ s_i &\geqq 0 \quad \text{for} \quad i \in M \setminus I. \end{aligned}$$

When we come to Step 8, we have only to increase the slack variable s_k to find a vertex v.

(v) We have assumed that system (4.1) is nondegenerate. Though it is degenerate, we can obtain the same result theoretically by perturbing the right side constant $-c \in R^n$ by $(\varepsilon, \varepsilon^2, \ldots \varepsilon^n)^t \in R^n$ for a sufficiently small ε, and practically by augmenting system (4.1) to

$$AX = \left[\begin{array}{c|c} 1 & 0 \\ \hline -c & I_n \end{array}\right]$$

and considering lexico-positive solution $X \in R^{(|I|+|U|+1)\times(n+1)}$, where I_n is the $n \times n$ identity matrix and $A \in R^{(n+1)\times(|I|+|U|+1)}$ is the coefficient matrix representing (4.1). Note that the first column of X is the solution of (4.1) with the right side $-c$ and $X(1, \varepsilon, \varepsilon^2, \ldots, \varepsilon^n)^t$ is the solution of (4.1) with $-c + (\varepsilon, \varepsilon^2, \ldots, \varepsilon^n)^t$.

(vi) Let $(\lambda, \mu) \in A(I, U(I))$, $z = \Sigma_{u \in U(I)} \lambda_u u$. If $\lambda_w \neq 1$, then $z/(1-\lambda_w)$ is in $F(I)$ or equivalently

$$\begin{aligned} a^i \cdot z - (1-\lambda_w) b_i &\leqq 0 \quad \text{for} \quad i \in M \setminus I \\ &= 0 \quad \text{for} \quad i \in I. \end{aligned}$$

Therefore when $\lambda_w \neq 1$, (4.1) is equivalent to

$$\sum_{i\in I} \mu_i a^i + Cz + \lambda_w Cw = -c$$

$$0 \leqq \lambda_w \leqq 1$$

$$\mu_i \geqq 0 \quad \text{for} \quad i \in I$$

$$(5.1) \qquad a^i \cdot z + b_i \lambda_w + s_i = b_i \quad \text{for} \quad i \in M \setminus I$$

$$a^i \cdot z + b_i \lambda_w = b_i \quad \text{for} \quad i \in I$$

$$s_i \geqq 0 \quad \text{for} \quad i \in M \setminus I.$$

We could use system (5.1) to trace path S, which would release us from column generation. Note however that (5.1) has $m-1$ more equations than (4.1).

(vii) Let us consider the numerical example in Eaves [4]. The matrix and vectors defining the problem are

$$D = \begin{bmatrix} 1 & -2 \\ 1 & -1 \end{bmatrix} \qquad c = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

$$[a^1 \ a^2 \ a^3 \ a^4 \ a^5 \ a^6] = \begin{bmatrix} -1 & 1 & 0 & -1 & 0 \\ -1 & -1 & 1 & 0 & -1 \end{bmatrix}$$

$$[b_1 \ b_2 \ b_3 \ b_4 \ b_5 \ b_6] = [-1 \ \ 0 \ \ 4 \ \ 0 \ \ 0] \ .$$

The stationary point is $\hat{x} = (4,4)^t$, which is a vertex of Ω having the dual cone

$$F^* = \{ \ (y_1, y_2) \ : \ y_1 \geqq 0, \quad y_1 + y_2 \geqq 0 \ \}.$$

It is easy to see that $-f(x)$ lies in this dual cone for any point of Ω (check the function value at each vertex, for instance). This means that the function value at any point of Ω suggests us that $\hat{x}$ is a plausible candidate for a stationary point. In fact, whichever point of Ω we may choose as the starting point, the algorithm yields $\hat{x}$ by only one pivot operation after it solves the linear programming problem in Step 1. This example suggests us that the algorithm is promising when we have a good guess of a stationary point. When a series of gradually changing stationary point problems is being solved, the algorithm will be efficient because the stationary point of the former problem usually serves as a good starting point of the successive problem.

(viii) Though we have made no assumption on the function $f(x) = Dx + c$, we have assumed the compactness of Ω, which guarantees both the existence of a stationary point and the finiteness of the algorithm. It will need further work to extend the algorithm for problems with unbounded polyhedral sets Ω.

Acknowledgment

This work was carried out while the author was supported by the Alexander von Humboldt-Foundation, West Germany, and was visiting the Institute of Econometrics and Operations Research, the University of Bonn. He wishes to thank the foundation and the institute for their friendly support. He also wishes to thank Dolf Talman, Tilburg University, for inspiring discussion and referees for their helpful comments.

References

[1] Dantzig, G.B., Linear Programming and Extensions (Princeton University Press, Princeton, 1963).

[2] Eaves, B.C., "A short course in solving equations with PL homotopies", SIAM-AMS Proceedings 9 (1978) 73-143.

[3] Eaves, B.C., "Computing stationary points", Mathematical Programming Study 7 (1978) 1-14.

[4] Eaves, B.C., "Computing stationary point, again", in: O.L. Mangasarian et al. eds., Nonlinear Programming 3 (Academic Press, New York, 1978) pp.391-405.

[5] Garcia, C.B. and W.I. Zangwill, Pathways to Solutions, Fixed Points and Equilibria (Prentice-Hall, Englewood Cliffs, 1981).

[6] Van der Heyden, L., "A variable dimension algorithm for linear complementarity problem", Mathematical Programming 19 (1980) 328-346.

[7] Kojima, M. and Y. Yamamoto, "Variable dimension algorithms: basic theory, interpretations and extensions of some existing methods", Mathematical Programming 24 (1982) 177-215.

[8] van der Laan, G. and A.J.J. Talman, "An algorithm for the linear complementarity problem with upper and lower bounds", FEW 200, Department of Economics, Tilburg University (Tilburg, the Netherlands, 1985).

[9] Lemke, C.E., "Bimatrix equilibrium points and mathematical programming", Management Sciences 11 (1965) 681-689.

[10] Pang, J.S., "A column generation technique for the computation of stationary points", Mathematics of Operations Research 6 (1981) 213-224.

[11] Reiser, P.M., "A modified integer labelling for complementarity problems", Mathematics of Operations Research 6 (1981) 129-139.

[12] Talman, A.J.J. and L. Van der Heyden, "Algorithms for the linear complementarity problem wich allow an arbitrary starting point", in: B.C. Eaves et al. eds., Homotopy Methods and Global Convergence (Plenum Press, New York, 1983) pp.267-285.

[13] Talman, A.J.J. and Y. Yamamoto, "A globally convergent simplicial algorithm for stationary point problems on polytopes", Report No.86422-OR, Institut für Ökonometrie und Operations Research, Universität Bonn (Bonn, West Germany, 1986).

Yoshitsugu Yamamoto
Institute of Socio-Economic Planning
University of Tsukuba
Sakura, Ibaraki 305, Japan
Tel. 0298-53-5369