AN IMPLEMENTABLE ACTIVE-SET ALGORITHM FOR COMPUTING A B-STATIONARY POINT OF A MATHEMATICAL PROGRAM WITH LINEAR COMPLEMENTARITY CONSTRAINTS: ERRATUM*

MASAO FUKUSHIMA† AND PAUL TSENG‡

Abstract. In [3], an ϵ-active set algorithm was proposed for solving a mathematical program with a smooth objective function and linear inequality/complementarity constraints. It is asserted therein that, under a uniform LICQ on the ϵ-feasible set, this algorithm generates iterates whose cluster points are B-stationary points of the problem. However, the proof has a gap and only shows that each cluster point is an M-stationary point. We discuss this gap and show that B-stationarity can be achieved if the algorithm is modified and an additional error bound condition holds.

Key words. MPEC, B-stationary point, ϵ-active set, error bound

AMS subject classifications. 65K05, 90C30, 90C33

1. Introduction. In a recent paper by the authors [3], an ϵ-active set algorithm was proposed for solving the following mathematical program with equilibrium constraints (MPEC):

$$\begin{array}{llll} & \text{minimize} & f(z) & \\ & \text{subject to} & G_i(z) \geq 0, & i = 1, \ldots, m, \\ & & H_i(z) \geq 0, & i = 1, \ldots, m, \\ (1) & & G_i(z) H_i(z) = 0, & i = 1, \ldots, m, \\ & & g_j(z) \leq 0, & j = 1, \ldots, p, \\ & & h_l(z) = 0, & l = 1, \ldots, q, \end{array}$$

where f is a real-valued continuously differentiable function on $\Re^n$ and G_i, H_i, g_j, h_l are real-valued *affine* functions on $\Re^n$. In Theorem 4.1(a) of [3], it is asserted that every cluster point of iterates generated by the algorithm is a B-stationary point of (1). However, the proof has a gap and only shows that every cluster point is an M-stationary point. We will discuss this gap and a modified algorithm that achieves B-stationarity under an additional error bound condition.

The gap occurs on [3, page 734] in the line "If $\nu_k \to 0$, then $|\mathcal{K}'| = \infty$, $\delta_k \to 0$, and the updating formula for ϵ_k would imply $\epsilon_k \to 0$, so any cluster point $\bar{z}$ of $\{\hat{z}^k\}_{k \in \mathcal{K}'}$ would be a KKT point of the *relaxed problem* $\mathrm{R}(\bar{z})$, which is a B-stationary point of MPEC (1) under the uniform LICQ." In particular, we have for all $k \in \mathcal{K}'$ that

$$(2) \qquad v_i^k \geq -\nu_k \quad \text{and} \quad w_i^k \geq -\nu_k \quad \forall\, i \in \hat{A}^k \cap \hat{B}^k,$$

where $\hat{A}^k$, $\hat{B}^k$ are given by [3, Eq. (7)] and v_i^k, w_i^k are multipliers associated with $\hat{z}^k$ (see [3, Eqs. (5), (6)]).[1] Thus, if a subsequence $\{\hat{z}^k\}_{k \in \mathcal{K}''}$ ($\mathcal{K}'' \subseteq \mathcal{K}'$) converges to

*This research is supported by Scientific Research Grant-in-Aid from the Ministry of Education, Science, Sports and Culture of Japan, and by National Science Foundation grant DMS-0511283.

†Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan (fuku@amp.i.kyoto-u.ac.jp).

‡Department of Mathematics, University of Washington, Seattle, WA 98195 (tseng@math.washington.edu).

[1] Throughout, we use the same notations as [3].

some $\bar{z}$, then by further passing to a subsequence if necessary, we can assume that the index sets $\hat{A}^k$ and $\hat{B}^k$ are constant (i.e., $\hat{A}^k = \bar{A}$, $\hat{B}^k = \bar{B}$ for some $\bar{A}$, $\bar{B}$) for all $k \in \mathcal{K}''$. Since $\bar{z}$ satisfies the uniform LICQ, $\{(v_i^k)_{i\in\bar{A}}, (w_i^k)_{i\in\bar{B}}\}_{k\in\mathcal{K}''}$ also converges to some $(\bar{v}_i)_{i\in\bar{A}}, (\bar{w}_i)_{i\in\bar{B}}$.[2] By (2),

$$\bar{v}_i \geq 0 \quad \text{and} \quad \bar{w}_i \geq 0 \quad \forall\, i \in \bar{A} \cap \bar{B}.$$

This together with [3, Eqs. (5), (6)] implies that $\bar{z}$ is an *M-stationary point* (see [4, 5] and (5)). If in addition

$$\bar{A} \cap \bar{B} = A_0(\bar{z}) \cap B_0(\bar{z}), \tag{3}$$

then $\bar{z}$ is a B-stationary point of (1). In general, however, we can only assert that $\bar{A} \cap \bar{B} \subseteq A_0(\bar{z}) \cap B_0(\bar{z})$. This is the gap.

2. A modified ϵ-active set algorithm. We now describe a way, based on an active set identification approach of Facchinei, Fischer, and Kanzow [1], to modify the ϵ-active set algorithm so that (3) holds under an additional error bound condition. To simplify the notation, we will consider only the complementarity constraints, i.e., we assume $p = q = 0$ in (1). The general case can be treated analogously. The Lagrangian associated with (1) is

$$L(z, v, w) := f(z) + \sum_{i=1}^{m} (G_i(z)v_i + H_i(z)w_i).$$

We assume that there exists a computable continuous function $R : \Re^n \times \Re^m \times \Re^m \to [0, \infty)$ providing a local Hölder error bound at each M-stationary point $\bar{z}$, i.e., there exist scalars $\tau > 0$, $\gamma > 0$, and $\delta > 0$ (depending on $\bar{z}$) such that

$$\|(z, v, w) - (\bar{z}, \bar{v}, \bar{w})\| \leq \tau R(z, v, w)^\gamma \quad \text{whenever } \|(z, v, w) - (\bar{z}, \bar{v}, \bar{w})\| \leq \delta, \tag{4}$$

where the multiplier vectors $\bar{v}, \bar{w}$ satisfy

$$\nabla_z L(\bar{z}, \bar{v}, \bar{w}) = 0, \ \left\{ \begin{array}{c} \bar{v}_i \perp G_i(\bar{z}) \geq 0 \\ \bar{w}_i \perp H_i(\bar{z}) \geq 0 \end{array} \right\}, \ G_i(\bar{z})H_i(\bar{z}) = 0, \ \left\{ \begin{array}{c} \bar{v}_i\bar{w}_i \geq 0, \\ \bar{v}_i \geq 0 \text{ or } \bar{w}_i \geq 0 \end{array} \right\} \ \forall i. \tag{5}$$

Here, $a \perp b$ means $ab = 0$. Due to uniform LICQ, $\bar{v}, \bar{w}$ are uniquely determined by $\bar{z}$. In fact, (5) characterizes M-stationarity for any $\bar{z} \in \Re^n$. We also assume that

$$R(\bar{z}, \bar{v}, \bar{w}) = 0 \quad \Longleftrightarrow \quad (\bar{z}, \bar{v}, \bar{w}) \text{ satisfies (5).} \tag{6}$$

The "residual" function $R(z, v, w)$ can be constructed analogous to the NLP and NCP cases [1, 2]. In particular, consider

$$\begin{aligned} R(z, v, w) \ &:= \ \|\nabla_z L(z, v, w)\| + \sum_{i=1}^{m} \Big(|\min\{G_i(z), |v_i|\}| + |\min\{H_i(z), |w_i|\}| \\ &+ \ |G_i(z)H_i(z)| + |\min\{0, v_iw_i\}| + |\min\{0, v_i\}\min\{0, w_i\}| \Big). \end{aligned} \tag{7}$$

[2]This follows from [3, Eq. (6)], $\|r^k\|_1 \leq \delta_k \to 0$ (see [3, Eq. (5)]), and the fact that if $b^k = C^k u^k$ for all k and $b^k \to b \in \Re^q$, $C^k \to C \in \Re^{q\times p}$ with C having linearly independent columns, then $u^k \to u \in \Re^p$ with u being the unique solution of $b = Cu$.

Then, R is continuous and satisfies (6). Arguing as in the proof of Cor. 6.6.4 in [2], we have that the local error bound (4) holds if the M-stationary point $\bar{z}$ is isolated and f and ∇f are continuous and subanalytic (G and H, by being affine, are automatically continuous and subanalytic). A referee suggests that the assumption of $\bar{z}$ being isolated is benign when G and H are affine. In particular, it is readily shown that the M-stationary points of (1) are isolated if f is strictly convex on the null space of the active constraint gradients. Alternatively, it can be shown that the local error bound (4) holds with $\gamma = 1$ if a certain 2nd-order sufficient condition holds at $\bar{z}$. This is a topic for further research.

Using (4), (6) and following [1, 2], the function

$$\Theta(z, v, w) := -1/\log(\min\{R(z, v, w), 0.9\})$$

has the active set identification property that, for any M-stationary point $\bar{z}$ and corresponding multiplier vectors $\bar{v}, \bar{w}$, we have

$$\lim_{(z,v,w)\to(\bar{z},\bar{v},\bar{w})} \frac{G_i(z)}{\Theta(z, v, w)} = \begin{cases} 0 & \text{if } G_i(\bar{z}) = 0, \\ \infty & \text{if } G_i(\bar{z}) > 0, \end{cases}$$

and similarly with "G_i" replaced by "H_i". Let us define

$$\bar{A}^k := \left\{ i \in \{1, ..., m\} : \frac{G_i(\hat{z}^k)}{\Theta(\hat{z}^k, v^k, w^k)} \leq 1 \right\},$$
$$\bar{B}^k := \left\{ i \in \{1, ..., m\} : \frac{H_i(\hat{z}^k)}{\Theta(\hat{z}^k, v^k, w^k)} \leq 1 \right\},$$

where the ith component of v^k is v_i^k if $i \in \hat{A}^k$ and is zero otherwise (and w^k is defined analogously). Since $(\hat{z}^k, v^k, w^k)$ satisfies [3, Eqs. (4)-(6)], if (2) holds, then $R(\hat{z}^k, v^k, w^k)$ would tend to zero as $\hat{z}^k \to \bar{z}$ and $\epsilon_k, \delta_k, \nu_k$ tend to zero and, for $\hat{z}^k$ sufficiently near $\bar{z}$, we would have (v^k, w^k) sufficiently near $(\bar{v}, \bar{w})$ (due to [3, A2]) and

(8) $$\bar{A}^k = A_0(\bar{z}), \quad \bar{B}^k = B_0(\bar{z}),$$

as well as

(9) $$A_\epsilon(\hat{z}^k) \supseteq \bar{A}^k \supseteq \hat{A}^k, \qquad B_\epsilon(\hat{z}^k) \supseteq \bar{B}^k \supseteq \hat{B}^k,$$

where $\epsilon \geq 0$ is defined as in [3] (see page 727 therein).[3] Let

(10) $$\bar{\epsilon}_k := \max\left\{ \epsilon_k, \max_{i\in\bar{A}^k} G_i(\hat{z}^k), \max_{i\in\bar{B}^k} H_i(\hat{z}^k) \right\}.$$

Since $\bar{\epsilon}_k \geq \epsilon_k$, [3, Eq. (4)] implies that $\hat{z}^k \in \mathcal{F}_{\bar{\epsilon}_k}[A^k, B^k]$ for all k. In fact, it can be seen that $\hat{z}^k$ remains an approximate KKT point of the subproblem [3, Eq. (3)] (in the sense of [3, Eqs. (4)-(6)]) when ϵ_k is replaced by $\bar{\epsilon}_k$ and $\hat{A}^k, \hat{B}^k$ are correspondingly replaced by $A_{\bar{\epsilon}_k}(\hat{z}^k), B_{\bar{\epsilon}_k}(\hat{z}^k)$. Thus, we can modify Step 2 of the ϵ-active set algorithm by possibly making this replacement when we are in case (c) and (9) holds.

THE MODIFIED $\boldsymbol{\epsilon}$-ACTIVE SET ALGORITHM FOR MPEC (1).

[3]The first containment in (9) holds whenever $\Theta(\hat{z}^k, v^k, w^k) \leq \epsilon$, which in turn holds whenever $R(\hat{z}^k, v^k, w^k)$ is sufficiently small. By (8) and [3, Eq. (7)], the second containment in (9) holds whenever $A_0(\bar{z}) \supseteq A_{\epsilon_k}(\hat{z}^k)$, which in turn holds whenever $\hat{z}^k$ is near $\bar{z}$ and ϵ_k is sufficiently small. The other two containments can be argued similarly.

This is the same as the ϵ-active set algorithm in [3, pp. 730-731], except that when we are in case (c) in Step 2, we do the following: If

$$\text{(9) holds,} \quad \bar{A}^k \cap \bar{B}^k \neq \hat{A}^k \cap \hat{B}^k, \quad \bar{\epsilon}_k < \bar{\epsilon} \tag{11}$$

($\bar{\epsilon}$ is a threshold which initially can be any positive scalar below ϵ), then repeat Step 2 with ϵ_k replaced by $\bar{\epsilon}_k$ (and with $\hat{A}^k$, $\hat{B}^k$ redefined accordingly, i.e., they are replaced by $A_{\bar{\epsilon}_k}(\hat{z}^k), B_{\bar{\epsilon}_k}(\hat{z}^k)$ in Step 2, (9), (11)), and update $\bar{\epsilon} \leftarrow \bar{\epsilon}/2$. Otherwise, if $\epsilon_k \leq \epsilon_{\text{tol}}$ and $\nu_k \leq \nu_{\text{tol}}$, then terminate; otherwise, determine ν_{k+1} and $\tilde{z}^k$ by [3, Eq. (14)], and proceed to Step 3.

If (11) holds, then $\epsilon_k < \bar{\epsilon}_k$,[4] which in turn implies $\bar{A}^k = A_{\bar{\epsilon}_k}(\hat{z}^k)$ and $\bar{B}^k = B_{\bar{\epsilon}_k}(\hat{z}^k)$.[5] Thus, when Step 2 is repeated, the second relation in (11) is violated.

THEOREM 2.1. *Under assumptions [3, A1–A3], the following results hold for the sequence* $\{(z^k, \hat{z}^k, \tilde{z}^k, \epsilon_k, \nu_k)\}$ *generated by the modified ϵ-active set algorithm, with* $\bar{\mathcal{K}} := \{k :$ *at iteration k, Step 2 is repeated*$\}$.

(a) *Suppose that each M-stationary point $\bar{z}$ of MPEC* (1) *that is not B-stationary satisfies* (4)*, where $(\bar{v}, \bar{w})$ satisfies* (5) *and R satisfies* (6). *If $\epsilon_0 > 0$, $\nu_0 > 0$, $\epsilon_{\text{tol}} = \nu_{\text{tol}} = 0$, f is Lipschitz continuous with constant L on a set Z containing $\{z^k\}$ and $\{\tilde{z}^k\}$, and $|\bar{\mathcal{K}}| < \infty$ (respectively, $|\bar{\mathcal{K}}| = \infty$), then $\epsilon_k \downarrow 0$, $\nu_k \downarrow 0$, and every cluster point of $\{\hat{z}^k\}$ (respectively, $\{\hat{z}^k\}_{k \in \bar{\mathcal{K}}}$) is a B-stationary point of MPEC* (1).

(b) *If $\epsilon_0 = \nu_0 = 0$ and f is quadratic, then there exists a $\bar{k} \in \{0, 1, \ldots\}$ such that $\hat{z}^{\bar{k}}$ is a B-stationary point of MPEC* (1).

Proof. The first paragraph of the proof is identical to the proof of [3, Thm. 4.1], except we define $\mathcal{K} := \{k :$ We enter Step 3 from case (a) or (b) in Step 2 at iteration $k\}$ and $\mathcal{K}' := \{k :$ We enter Step 3 from case (c) in Step 2 at iteration $k\}$. The proof of (b) is identical to the proof of [3, Thm. 4.1(b)]. We prove (a) below.

(a) Suppose $\nu_k \to 0$. Then $|\mathcal{K}'| = \infty$, $\delta_k \to 0$, and the updating formula for ϵ_k and $\bar{\epsilon}$ imply $\epsilon_k \to 0$, so any cluster point $\bar{z}$ of $\{\hat{z}^k\}_{k \in \mathcal{K}'}$ is an M-stationary point of MPEC (1). First, suppose $|\bar{\mathcal{K}}| < \infty$, so that $\bar{\epsilon} > 0$ is constant after a while. Let $\{\hat{z}^k\}_{k \in \mathcal{K}''}$ ($\mathcal{K}'' \subseteq \mathcal{K}'$) be any subsequence converging to $\bar{z}$. Since [3, Eqs. (4)-(6)] and (2) hold for all $k \in \mathcal{K}''$, we have from [3, A2] and the same argument as in Section 1 that $\{(v^k, w^k)\}_{k \in \mathcal{K}''} \to (\bar{v}, \bar{w})$ satisfying (5). By (6), $R(\bar{z}, \bar{v}, \bar{w}) = 0$. Since R is continuous, $\{R(\hat{z}^k, v^k, w^k)\}_{k \in \mathcal{K}''} \to 0$. If $\bar{z}$ is not B-stationary for (1), then the error bound (4) would hold and this would imply that (8) and (9) hold for all $k \in \mathcal{K}''$ sufficiently large. Moreover, $\{\bar{\epsilon}_k\}_{k \in \mathcal{K}''} \to 0$, so that $\bar{\epsilon}_k < \bar{\epsilon}$ for all $k \in \mathcal{K}''$ sufficiently large. Thus, at each such iteration $k \in \mathcal{K}''$, we would have upon entering Step 3 that $\bar{A}^k \cap \bar{B}^k = \hat{A}^k \cap \hat{B}^k$ (since (11) must be violated). Then it would follow from (2) and (8) that $\bar{z}$ is a B-stationary point of (1), a contraction. Second, suppose $|\bar{\mathcal{K}}| = \infty$. Then, as we discussed earlier, for each iteration $k \in \bar{\mathcal{K}}$, the second relation in (11) is violated upon entering Step 3, i.e., $\bar{A}^k \cap \bar{B}^k = \hat{A}^k \cap \hat{B}^k$. Then, an argument similar to the one above shows that every cluster point $\bar{z}$ of $\{\hat{z}^k\}_{k \in \bar{\mathcal{K}}}$ is a B-stationary point of (1).

[4] If $\epsilon_k = \bar{\epsilon}_k$, then (10) and [3, Eq. (7)] would imply $\bar{A}^k \subseteq \hat{A}^k$ and $\bar{B}^k \subseteq \hat{B}^k$, so (9) would yield $\bar{A}^k = \hat{A}^k$ and $\bar{B}^k = \hat{B}^k$, contradicting (11).

[5] Why? Since $\epsilon_k < \bar{\epsilon}_k$, we have from (10) and the definition of $\bar{A}^k$ and $\bar{B}^k$ that

$$\bar{\epsilon}_k = \max\left\{ \max_{i \in \bar{A}^k} G_i(\hat{z}^k), \max_{i \in \bar{B}^k} H_i(\hat{z}^k) \right\} \leq \Theta(\hat{z}^k, v^k, w^k).$$

Thus, if $i \notin \bar{A}^k$, then $G_i(\hat{z}^k) > \Theta(\hat{z}^k, v^k, w^k) \geq \bar{\epsilon}_k$. By (10), if $i \in \bar{A}^k$, then $G_i(\hat{z}^k) \leq \bar{\epsilon}_k$. This shows that $\bar{A}^k = A_{\bar{\epsilon}_k}(\hat{z}^k)$. An analogous argument shows that $\bar{B}^k = B_{\bar{\epsilon}_k}(\hat{z}^k)$.

Suppose instead $\nu_k \not\to 0$, so that $|\mathcal{K}'| < \infty$, $|\mathcal{K}| = \infty$, and $\nu = \lim_{k\to\infty} \nu_k > 0$. The remainder of the proof is identical to the proof of [3, Thm. 4.1(a)], except that, due to ϵ_k being replaced by $\bar{\epsilon}_k$ in Step 2 for all iterations $k \in \bar{\mathcal{K}}$, instead of [3, Eq. (22)] we have

$$f(z^{k+1}) \leq f(\tilde{z}^k) + 2L\tau m(\epsilon_k - \epsilon_{k+1} + \Delta_k) \quad \forall\, k,$$

where $\Delta_k := \bar{\epsilon}_k$ if $k \in \bar{\mathcal{K}}$ and $\Delta_k := 0$ otherwise. Since (11) holds at each iteration $k \in \bar{\mathcal{K}}$ and $\bar{\epsilon}$ is halved at each such iteration, it follows that $\sum_{k=0}^{\infty} \Delta_k = \sum_{k\in\bar{\mathcal{K}}} \bar{\epsilon}_k < \infty$. Then it can be argued similarly as in the proof of [3, Thm. 4.1(a)] that $\{f(z^k)\}$ converges and so on. □

We illustrate the assumptions of Theorem 2.1 with the following example of (1):

$$\text{minimize} \quad f(z) \quad \text{subject to} \quad z_1 \geq 0,\ z_2 \geq 0,\ z_1 z_2 = 0.$$

This example satisfies assumption [3, A2] for any $\epsilon \geq 0$. If $f(z) = (z_2)^p$ $(p \geq 1)$, then assumption [3, A1] also holds and each M-stationary point, which is of the form $(\bar{z}_1, 0)$ with $\bar{z}_1 \geq 0$, is B-stationary. If $f(z) = z_1^4 + z_2^2 - z_2$, then assumptions [3, A1, A3] also hold and the M-stationary points, $(0,0)$ and $(0, \frac{1}{2})$, are isolated with $(0, \frac{1}{2})$ B-stationary. For R given by (7), the error bound (4) holds at $(0,0)$. However, if $f(z) = z_2^2 - z_2$, then the M-stationary point $\bar{z} = (0,0)$, with multipliers $\bar{v} = 0, \bar{w} = -1$, is not B-stationary and is not isolated. Moreover, for any continuous R satisfying (6), the error bound (4) does not hold at $(0,0)$. This is because, for any fixed $\delta > 0$, $(\delta, 0)$ is M-stationary with multipliers $v = 0, w = -1$, so $R((\delta, x_2), 0, -1) \to R(\delta, 0), 0, -1) = 0$ as $x_2 \to 0$. But $\|((\delta, x_2), 0, -1) - ((0,0), 0, -1)\| \to \delta$ as $x_2 \to 0$.

Acknowledgments. The authors thank Lifeng Chen for notifying them of the gap in the proof of [3, Thm. 4.1]. They also thank two referees for their helpful comments.

REFERENCES

[1] F. Facchinei, A. Fischer, and C. Kanzow, *On the accurate identification of active constraints*, SIAM J. Optim., 9 (1999), pp. 14-32.

[2] F. Facchinei, and J.-S. Pang, Finite-Dimensional Variational Inequalities and Complementarity Problems, Vol. II, Springer-Verlag, New York, 2003.

[3] M. Fukushima and P. Tseng, *An implementable active-set algorithm for computing a B-stationary point of a mathematical program with linear complementarity constraints*, SIAM J. Optim., 12 (2002), pp. 724–739.

[4] J. Outrata, *Optimality conditions for a class of mathematical programs with equilibrium constraints*, Math. Oper. Res., 24 (1999), pp. 627–644.

[5] H. Scheel and S. Scholtes, *Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity*, Math. Oper. Res., 25 (2000), pp. 1–22.