New active set identification for general constrained optimization and minimax problems ☆

Yi Liu[a,b], Jin-bao Jian[a,*]

[a] *School of Mathematics and Information Science, Yulin Normal University, Yulin 537000, China*
[b] *College of Mathematics and Information Science as well as XingJian College of Science and Liberal art , Guangxi University, Nanning 530004, China*

Abstract

The purpose of this paper is to discuss the problem of identifying the active constraints for general constrained nonlinear programming and constrained minimax problems at an isolated local solution. Facchinei et al. [F. Facchinei, A. Fischer, and C. Kanzow, On the accurate identification of active constraints, SIAM J. Optim., 9(1998), 14-32] proposed an effective technique which can identify the active set in a neighborhood of an isolated local solution for nonlinear programming. Han et al. [D. L. Han, J. B. Jian and J. Li, On the accurate identification of active set for constrained minimax problems, Nonl. Anal.–TMA, 74(2011) 3022-3032] improved and extended it to minimax problems. In this work, a new active constraint identification set is constructed, not only is it tighter than the previous two identification sets, but also it can be used effectively in penalty algorithms. Without strict complementarity and linear independence, it is shown that the new identification technique can accurately identify the active constraints of nonlinear programming and constrained minimax problems. Some numerical examples illustrate the identification technique.

Keywords: constrained optimization; minimax problems; active constraints; identification technique.
AMS Subject Classification: 90C30, 65K10

1. Introduction

The correct identification of active constraints in constrained nonlinear programming and in constrained minimax problems is a very interesting problem, since it is important from both the theoretical and practical points of view for optimization problems. On the one hand, it can improve the local convergence behavior of algorithms, since the study of the local convergence rate of most algorithms implicitly or explicitly depends on the fact that the active set is eventually identified. On the other hand, it can considerably simplify algorithms and decrease the computation cost, as it removes the combinatorial aspect of the problem and locally reduces the inequality constrained minimization problem to an equality constrained problem which is easy to deal with.

Some earlier studies of the active set identification technique can be found in Refs. [1, 2, 3, 4], Bertsekas [1] proposed a two-metric algorithm for minimizing a nonlinear function subject to bound constraints on the components of x. A key aspect of this method is estimation of the active bounds at the solution. The similar result of [1] is also proved by Lescrenier [4] for a trust-region algorithm. Burke and Moré [2] took a geometric approach, the constraints need to be expressed in the form $x \in \Omega$, where Ω is a convex set. The set Ω can be partitioned into faces, where active set identification corresponds to the identification of the

☆Project supported by NSFC (Grant Nos. 11271086,11171250), and the Natural Science Foundation of Guangxi Province (Grant No. 2011GXNSFD018022) as well as Innovation Group of Talents Highland of Guangxi higher School.

*Corresponding author: jianjb@gxu.edu.cn or jianjb@gmail.com (J.-B. Jian), URL: http://jians.gxu.edu.cn.

face that contains the solution x^*. Burke [3] took a partly algebraic viewpoint and further developed the work in [2].

Wright [5] used a hybrid geometric-algebraic viewpoint and considered convex constraint sets Ω with boundaries defined by inequalities. Hare and Lewis [6] extended it to nonconvex sets. Active set identification for nonsmooth functions was developed by Lewis [7] and others. At the same time, the active set identification technique was extended to linear complementarity problems [8, 9, 10, 11].

Facchinei Fischer and Kanzow [12] proposed a technique based on the algebraic representation of the constraint sets. The work introduced an identification technique which can identify all the active constraints in a neighborhood of an isolated local solution and requires neither strict complementarity nor uniqueness of the multipliers. Oberlin and Wright [13] focused on identification schemes that do not require good estimates of the Lagrange multipliers to be available a priori. Further associated research can be seen in [14]

Jian [15, Section 1.5] summarized and improved the identification technique in [12, 13] such that it is tighter and suitable for infeasible algorithms, e.g., the strongly sub-feasible directions methods such as [16, 17]. Han, Jian and Li [18] developed the identification technique [15] to constrained minimax problems.

In this paper, motivated by the active set identification technique in Refs. [12, 15, 18] and the need for penalty algorithms, a new active set identification for general constrained optimization and minimax problems is proposed. The identification set not only tighter than all the three identification sets in [12, 15, 18], but also can be used to improve the penalty function methods, e.g. Ref. [19].

The following notation will be employed in this paper. The Euclidean distance of a point y from a nonempty set S is abbreviated by $\text{dist}[y, S]$. The symbol $\|\cdot\|$ represents the Euclidean vector norm. The symbol $\mathcal{B}_\epsilon$ denotes the open Euclidean ball with radius $\epsilon > 0$ and center at the origin, the dimensions of $\mathcal{B}_\epsilon$ are clear from the context. For two vectors a and b, the notation $a \perp b$ denotes $a^T b = 0$. For any index set L, we denote vector function $f_L(x) = (f_j(x), j \in L)$, and vector $\lambda_L = (\lambda_j, j \in L)$, and the cardinality of any finite set L by $|L|$.

2. New active set identification for general constrained optimization

We consider the following smooth constrained optimization problem:

$$\min\{f_0(x): \ x \in \mathbb{R}^n \text{ such that } f_j(x) \le 0, \ j \in I \text{ and } f_j(x) = 0, \ j \in E\}, \tag{2.1}$$

where sets $I = \{1, 2, \ldots, m\}$ and $E = \{m+1, \ldots, m+l\}$, the functions $f_0, f_j \ (j \in I \cup E) : \mathbb{R}^n \to \mathbb{R}$ are all twice continuously differentiable. A vector $x^* \in \mathbb{R}^n$ is called a Karush-Kuhn-Tucker (KKT) point of (2.1), if there exists a multiplier vector $\lambda^* \in \mathbb{R}^{m+l}$ such that (x^*, λ^*) solves the KKT system:

$$g_0(x) + \sum_{j \in I \cup E} \lambda_j g_j(x) = 0, \quad 0 \le \lambda_I \perp f_I(x) \le 0, \quad f_E(x) = 0, \tag{2.2}$$

where gradient function $g_j(x) := \nabla f_j(x)$ for $j = 0, \ 1, \ \ldots, \ m+l$. A vector (x^*, λ^*) is said to be a KKT pair of (2.1) if it satisfies system (2.2).

For the sake of simplicity, for $x \in \mathbb{R}^n$ and a KKT point x^* of problem (2.1), we denote:

$$\begin{cases} I(x) := \{j \in I: \ f_j(x) = 0\}, \ \varphi(x) := \max\{0, f_j(x), j \in I\}, \\ I^-(x) := \{j \in I: \ f_j(x) \le 0\}, \ I^+(x) = \{j \in I: \ f_j(x) > 0\}, \\ \overline{f}_j(x) = \begin{cases} f_j(x), & \text{if } j \in I^-(x); \\ f_j(x) - \varphi(x), & \text{if } j \in I^+(x), \end{cases} \\ \Lambda(x^*) := \{\lambda: \ (x^*, \lambda) \text{ constitutes a KKT pair of } (2.1)\}, \ \mathcal{K}(x^*) := \{(x^*, \lambda): \ \lambda \in \Lambda(x^*)\}. \end{cases} \tag{2.3}$$

In what follows, x^* always denotes a fixed, isolated KKT point of (2.1), so there is a neighborhood of x^* which does not contain any other KKT point.

The following definition of identification function was introduced by [12], and it is very important to our theoretical analysis.

Definition 2.1. *A function $\rho(x,\lambda):\mathbb{R}^n\times\mathbb{R}^{m+l}\to[0,+\infty)$ is called an identification function for $\mathcal{K}(x^*)$, if (a) ρ is continuous on $\mathcal{K}(x^*)$; (b) $(x^*,\lambda^*)\in\mathcal{K}(x^*)$ implies $\rho(x^*,\lambda^*)=0$; and (c) if (x^*,λ^*) belongs to $\mathcal{K}(x^*)$, then*

$$\lim_{\substack{(x,\lambda)\to(x^*,\lambda^*)\\(x,\lambda)\notin\mathcal{K}(x^*)}}\frac{\rho(x,\lambda)}{\mathrm{dist}[(x,\lambda),\mathcal{K}(x^*)]}=+\infty.$$

By a suitable identification function $\rho(x,\lambda)$, for example, formula (2.10), the identification set of active inequality constraints in [12] is defined by

$$I(x,\lambda):=\{j\in I:\ f_j(x)+\rho(x,\lambda)\ge 0\}. \tag{2.4}$$

Since the set of violation constraints $I^+(x)\subseteq I(x,\lambda)$, when (x,λ) is not sufficiently close to $\mathcal{K}(x^*)$, the accuracy of set $I(x,\lambda)$ as an estimate for $I(x^*)$ is poor. For this reason, Jian [15, Section 1.5], also Ref. [18] improved the technique and introduced a new identification as follows:

$$\begin{aligned}\hat{I}(x,\lambda)\ &:=\{j\in I:\ \bar{f}_j(x)+\rho(x,\lambda)\ge 0\}\\ &=\{j\in I^-(x):\ f_j(x)+\rho(x,\lambda)\ge 0\}\cup\{j\in I^+(x):\ f_j(x)-\varphi(x)+\rho(x,\lambda)\ge 0\}.\end{aligned} \tag{2.5}$$

It is known that, in a penalty method for inequality constrained optimization, the active constraints in set $\{j\in I:\ f_j(x)=\varphi(x)\}$ play a key role. Motivated by this, to propose an effective identification technique for penalty method, we construct a new active constraint identification set by:

$$\overline{I}(x,\lambda):=\{j\in I:\ f_j(x)-\varphi(x)+\rho(x,\lambda)\ge 0\}. \tag{2.6}$$

Obviously, $\overline{I}(x,\lambda)\subseteq\hat{I}(x,\lambda)\subseteq I(x,\lambda)$, so it is reasonable to say $\overline{I}(x,\lambda)$ is tighter than $\hat{I}(x,\lambda)$ and $I(x,\lambda)$ for identifying the active set $I(x^*)$. An important role of the new active set identification technique is that it can be used to design effective penalty methods since $\{j\in I:\ f_j(x)=\varphi(x)\}\subseteq\overline{I}(x,\lambda)$.

The following two theorems show that the new identification set (2.6) can exactly identify the active inequality constraints of the KKT point x^*.

Theorem 2.2. *Suppose that x^* is an isolated KKT point for (2.1). Let $\rho(x,\lambda)$ be an identification function for $\mathcal{K}(x^*)$. Then, for any given $\lambda^*\in\Lambda(x^*)$, there exists a constant $\epsilon=\epsilon(\lambda^*)>0$ such that*

$$\overline{I}(x,\lambda)\equiv I(x^*),\ \forall\ (x,\lambda)\in(x^*,\lambda^*)+\mathcal{B}_\epsilon. \tag{2.7}$$

Proof The function $\rho(x,\lambda)$ is an identification function, so $\rho(x,\lambda)$ satisfies the conditions in Definition 2.1. Since $f_j(x)$ is continuously differentiable, $f_j(x)$ and $\varphi(x)$ are all locally Lipschitz continuous (see [15, Theorem 1.1.6]). Hence exists a constant $L>0$ such that relations

$$-f_j(x)\le -f_j(x^*)+L\|x-x^*\|,\ \forall\ j\in I,\quad \varphi(x)\le\varphi(x^*)+L\|x-x^*\| \tag{2.8}$$

hold for all x sufficiently close to x^*.

We first show that relation $I(x^*)\subseteq\overline{I}(x,\lambda)$ holds. Let $j\in I(x^*)$, i.e., $f_j(x^*)=0$. In view of $\varphi(x^*)=0$, according to (2.8) and Definition 2.1(c), when $(x,\lambda)\notin\mathcal{K}(x^*)$ in a sufficiently small neighborhood of (x^*,λ^*), we have

$$-f_j(x)+\varphi(x)\le 2L\|x-x^*\|\le 2L\mathrm{dist}[(x,\lambda),\mathcal{K}(x^*)]\le\rho(x,\lambda).$$

These along with (2.6) show that $j\in\overline{I}(x,\lambda)$. When $(x,\lambda)\in\mathcal{K}(x^*)$ in a sufficiently small neighborhood of (x^*,λ^*), then $x=x^*$ and $(x=x^*,\lambda)$ is a KKT pair of (2.1). From Definition 2.1(b), one has $\rho(x,\lambda)=0$ and

$$-f_j(x)+\varphi(x)=-f_j(x^*)+\varphi(x^*)=0=\rho(x,\lambda),$$

which also implies that $j \in \overline{I}(x,\lambda)$. Therefore, one has proved $I(x^*) \subseteq \overline{I}(x,\lambda)$.

On the other hand, if $j \notin I(x^*)$, i.e., $f_j(x^*) < 0$, it follows, by condition (a) in Definition 2.1 and $\rho(x^*,\lambda^*) = \varphi(x^*) = 0$, that $f_j(x) - \varphi(x) + \rho(x,\lambda) < 0$ if (x,λ) is sufficiently close to (x^*,λ^*), so $j \notin \overline{I}(x,\lambda)$. Therefore, $\overline{I}(x,\lambda) \subseteq I(x^*)$ holds. Hence, the proof is complete. □

Theorem 2.3. *Suppose that x^* is an isolated KKT point for (2.1). Let $\rho(x,\lambda)$ be an identification function for $\mathcal{K}(x^*)$. If the Mangasarian-Fromovitz constraint qualification (MFCQ) holds at x^*, i.e., vectors $\{g_j(x^*),\ j \in E\}$ are linearly independent and set*

$$D(x^*) := \{d \in \mathbb{R}^n :\ g_j(x^*)^T d < 0,\ \forall\ j \in I(x^*);\ g_j(x^*)^T d = 0,\ \forall\ j \in E\} \neq \emptyset,$$

then, there is a constant $\epsilon > 0$ such that

$$\overline{I}(x,\lambda) \equiv I(x^*),\ \forall\ (x,\lambda) \in \mathcal{K}(x^*) + \mathcal{B}_\epsilon. \tag{2.9}$$

Based on above Theorem 2.2, the proof of Theorem 2.3 is similar to [12, Theorem 2.3]. From the analysis above, we can see that the crucial point in the identification of active constraints is the identification function. We further discuss how to construct such a function for (2.1).

Similar to the technique in [12], for an estimate (x,λ) of a KKT pair (x^*,λ^*), we define map $\Phi(x,\lambda)$: $\mathbb{R}^n \times \mathbb{R}^{m+l} \to \mathbb{R}^n \times \mathbb{R}^{m+l}$ and identification function $\rho(x,\lambda) :\ \mathbb{R}^n \times \mathbb{R}^{m+l} \to [0,+\infty)$ as follows:

$$\Phi(x,\ \lambda) := \begin{pmatrix} \nabla_x L(x,\lambda) \\ \min\{-f_I(x),\ \lambda_I\} \\ f_E(x) \end{pmatrix},\ \rho(x,\lambda) := \|\Phi(x,\lambda)\|^\sigma, \tag{2.10}$$

where $\nabla_x L(x,\lambda) = g_0(x) + \sum_{j \in I \cup E} \lambda_j g_j(x)$, parameter σ such that $0 < \sigma < 1$.

Similar to the analysis of [18, Theorems 2.1, 2.2] and [12, Theorem 3.7], it is easy to get the following result.

Theorem 2.4. *Suppose that x^* is a KKT point of (2.1) and the MFCQ at x^* is satisfied. Assume that one of the following three second-order sufficient conditions (SOSC)* (a)–(c) *holds:*

(a) *there exists a constant $\gamma > 0$ such that*

$$d^T \nabla^2_{xx} L(x^*,\lambda^*) d \geq \gamma \|d\|^2,\ \forall\ \lambda^* \in \Lambda(x^*),\ \forall\ d \in M^0(x^*,\lambda^*),$$

with $I^+(x^*,\lambda^*) := \{j \in I :\ \lambda^*_j > 0\}$ and

$$M^0(x^*,\lambda^*) := \{d \in R^n :\ g_j(x^*)^T d = 0,\ \forall\ j \in I^+(x^*,\lambda^*) \cup E;\quad g_j(x^*)^T d \leq 0,\ \forall\ j \in I(x^*) \setminus I^+(x^*,\lambda^*)\};$$

(b) *it follows that*

$$d^T \nabla^2_{xx} L(x^*,\lambda^*) d > 0,\ \forall\ \lambda^* \in \Lambda(x^*),\ \forall\ d \in M^0(x^*,\lambda^*) \setminus \{0\};$$

(c) *set $\Lambda(x^*) = \{\lambda^*\}$ is single and*

$$d^T \nabla^2_{xx} L(x^*,\lambda^*) d > 0,\ \forall\ d \in M^0(x^*,\lambda^*) \setminus \{0\}.$$

Then, the function $\rho(x,\lambda)$ defined by (2.10) is an identification function for $\mathcal{K}(x^*)$. Furthermore, there exists an $\varepsilon > 0$ such that relation (2.9) holds.

Proof. First, by [15, Corollary 1.4.3], under the MFCQ and any one of the three conditions stated, one can conclude that x^* is an isolated KKT point of (2.1).

Second, we will show that conditions (a) and (b) are complete under the given assumptions. Obviously, condition (a) implies condition (b). Conversely, under condition (b), suppose by contradiction that condition (a) is'nt true. Then, for each $\frac{1}{k}$, there exist an $\beta^k \in \Lambda(x^*)$ and a $d^k \in M^0(x^*,\beta^k)$ such that

$$(d^k)^T \nabla^2_{xx} L(x^*,\beta^k) d^k < \frac{1}{k}\|d^k\|^2,\ \forall\ k = 1,\ 2,\cdots.$$

Let $\tilde{d}^k = \frac{\tilde{d}^k}{\|\tilde{d}^k\|}$. Then $\tilde{d}^k \in M^0(x^*, \beta^k)$, $\|\tilde{d}^k\| = 1$, and

$$(\tilde{d}^k)^T \nabla^2_{xx} L(x^*, \beta^k) \tilde{d}^k < \frac{1}{k}, \ \forall\, k = 1,\ 2, \cdots. \tag{2.11}$$

On the other hand, it is'nt difficult to know that $\Lambda(x^*)$ is compact by the MFCQ, so sequence $\{(\tilde{d}^k, \beta^k)\}$ is bounded. Therefore, there exists an infinite set K such that $(\tilde{d}^k, \beta^k) \xrightarrow{K} (\tilde{d}, \tilde{\beta})$, $\tilde{\beta} \in \Lambda(x^*)$, $\tilde{d} \neq 0$. Suppose, without loss of generality, that $I^+(x^*, \beta^k) \equiv \tilde{I}$, $\forall\, k \in K$. Taking into account $\tilde{d}^k \in M^0(x^*, \beta^k)$, one has

$$\nabla f_i(x^*)^{\mathrm{T}} \tilde{d}^k = 0,\ \forall\, i \in \tilde{I} \cup E;\ \nabla f_i(x^*)^{\mathrm{T}} \tilde{d}^k \le 0,\ \forall\, i \in I(x^*) \setminus \tilde{I},\ \forall\, k \in K.$$

Passing to the limit in the above relations for $k \to \infty$ and $k \in K$, we have

$$\nabla f_i(x^*)^{\mathrm{T}} \tilde{d} = 0,\ \forall\, i \in \tilde{I} \cup E;\ \nabla f_i(x^*)^{\mathrm{T}} \tilde{d} \le 0,\ \forall\, i \in I(x^*) \setminus \tilde{I}.$$

Again, in view of $I^+(x^*, \tilde{\beta}) \subseteq \tilde{I}$, the above relations shows that

$$\begin{gathered} \nabla f_i(x^*)^{\mathrm{T}} \tilde{d} = 0,\ \forall\, i \in I^+(x^*, \tilde{\beta}) \cup E \subseteq \tilde{I} \cup E, \\ \nabla f_i(x^*)^{\mathrm{T}} \tilde{d} \le 0,\ \forall\, i \in I(x^*) \setminus I^+(x^*, \tilde{\beta}) = (I(x^*) \setminus \tilde{I}) \cup (\tilde{I} \setminus I^+(x^*, \tilde{\beta})). \end{gathered}$$

Therefore, $\tilde{d} \in M^0(x^*, \tilde{\beta})$ and $\tilde{d} \neq 0$. However, passing to the limit in (2.11) for $k \to \infty$ and $k \in K$, it follows that $(\tilde{d})^{\mathrm{T}} \nabla^2_{xx} L(x^*, \tilde{\beta}) \tilde{d} \le 0$, and this contradicts condition (b). Hence, condition (b) implies condition (a), and conditions (a) and (b) are complete. Furthermore, the relations among the above three conditions are as follows: condition (c) $\Rightarrow$ condition (b) $\Rightarrow$ condition (a).

Third, referring to the analysis for [13, Theorems 2.1, 2.2] and [12, Theorem 3.7], under the MFCQ and condition (a), one can show that the function $\rho(x, \lambda)$ defined by (2.10) is an identification function. Therefore, $\rho(x, \lambda)$ defined by (2.10) is an identification function if any one of the three conditions holds.

Finally, by Theorem 2.3, there exists an $\varepsilon > 0$ such that relation (2.9) holds. So the proof is complete.□

3. Extension to minimax problems

In this section, we extend the proposed identification technique to the following constrained minimax problem:

$$\min\{F(x) :\ x \in \mathbb{R}^n \text{ such that } f_j(x) \le 0,\ j \in I \text{ and } f_j(x) = 0,\ j \in E\}, \tag{3.1}$$

where the function $F(x)$=$\max\{f_j(x), j \in J\}$. Assume that f_j $(j \in J \cup I \cup E) : \mathbb{R}^n \to \mathbb{R}$ are all twice continuously differentiable. We denote the active sets of (3.1) by

$$J(x) = \{j \in J :\ f_j(x) = F(x)\},\ I(x) = \{j \in I :\ f_j(x) = 0\},\ A(x) = J(x) \cup I(x).$$

A point x^* with multiplier $\lambda^* = (\lambda^*_j, j \in J \cup I \cup E)$ is said to be a stationary point to problem (3.1), if it satisfies:

$$\begin{cases} \sum\limits_{j \in J \cup I \cup E} \lambda^*_j g_j(x^*) = 0, \sum\limits_{j \in J} \lambda^*_j = 1, \\ 0 \le \lambda^*_J \perp (f_J(x^*) - F(x^*)e),\quad 0 \le \lambda^*_I \perp f_I(x^*) \le 0,\quad f_E(x^*) = 0, \end{cases} \tag{3.2}$$

And (x^*, λ^*) satisfying condition (3.2) is called a stationary pair of (3.1).

Since the objective function $F(x)$ is nondifferentiable, to extend the active set identifying technique in Section 2, by introducing an additional variable $z \in \mathbb{R}$, we reformulate the minimax problem (3.1) as the following smooth nonlinear program:

$$\min\{z :\ (x, z) \in \mathbb{R}^{n+1} \text{ such that } f_J(x) \le ze,\ f_I(x) \le 0,\ f_E(x) = 0\}. \tag{3.3}$$

Obviously, if (x^*, λ^*) is a stationary pair of (3.1), then (x^*, z^*, λ^*) with $z^* = F(x^*)$ is a KKT pair of (3.3). Furthermore, from [18, Proposition 3.1], one knows that the MFCQ and the (affine) LICQ (linearly independent constraint qualification) for problems (3.1) and (3.3) are respectively equivalent. For a given stationary point x^* to problem (3.1), we denote

$$\Lambda_{\rm mm}(x^*) := \{\lambda: \ (x^*, \lambda) \text{ is a stationary pair to (3.1)}\}, \ \mathcal{K}_{\rm mm}(x^*) := \{(x^*, \lambda): \ \lambda \in \Lambda_{\rm mm}(x^*)\},$$

where "mm" has the meaning "minimax". With the help of problem (3.3), in a similar fashion to [18, Section 3], we consider functions:

$$\Phi_{\rm mm}(x, \lambda) := \begin{pmatrix} \sum\limits_{j \in J \cup I \cup E} \lambda_j g_j(x) \\ \sum\limits_{j \in J} \lambda_j - 1 \\ \min\{F(x)e - f_J(x), \ \lambda_J\} \\ \min\{-f_I(x), \ \lambda_I\} \\ f_E(x) \end{pmatrix}, \ \rho_{\rm mm}(x, \lambda) := \|\Phi_{\rm mm}(x, \lambda)\|^{\sigma}. \tag{3.4}$$

Then, motivated by the proposed active set identifying technique in Section 2 and the active set identifying technique in [18, (3.5) and (3.6)], we consider the following active constraint identifying sets:

$$\begin{cases} J_{\rm mm}(x, \lambda) := \{j \in J: \ f_j(x) - F(x) + \rho_{\rm mm}(x, \lambda) \geq 0\}, \\ I_{\rm mm}(x, \lambda) := \{j \in I: \ f_j(x) - \varphi(x) + \rho_{\rm mm}(x, \lambda) \geq 0\}, \\ \overline{A}(x, \lambda) := J_{\rm mm}(x, \lambda) \cup I_{\rm mm}(x, \lambda). \end{cases} \tag{3.5}$$

By means of problem (3.3) and the discussions in Section 2 above, similar to the analysis in [18, Section 3], one has the following new result of accuracy identification for the minimax problem (3.1).

Theorem 3.1. *Suppose that is x^* is a stationary point of the minimax problem* (3.1)*, and assume that the MFCQ is satisfied at x^*. If one of the following three SOSC* (a)–(c) *holds:*

(a) *there exists a constant $\gamma > 0$ such that*

$$d^T \nabla^2_{xx} L_{mm}(x^*, \lambda^*) d \geq \gamma \|d\|^2, \ \forall \ \lambda^* \in \Lambda_{\rm mm}(x^*), \ \forall \ d \in M(x^*, \lambda^*),$$

where $\nabla^2_{xx} L_{\rm mm}(x^*, \lambda^*) = \sum\limits_{j \in J \cup I \cup E} \lambda^*_j \nabla^2 f_j(x^*)$*, and*

$$\begin{aligned} M(x^*, \lambda^*) = \quad & \{d \in R^n: \ g_j(x^*)^T d = 0, \ \forall \ j \in J^+(x^*, \lambda^*) \cup I^+(x^*, \lambda^*) \cup E; \\ & g_j(x^*)^T d \leq 0, \ \forall \ j \in (J(x^*) \cup I(x^*)) \setminus (J^+(x^*, \lambda^*) \cup I^+(x^*, \lambda^*))\}, \end{aligned}$$

$$J^+(x^*, \lambda^*) = \{j \in J: \ \lambda^*_j > 0\}, \ I^+(x^*, \lambda^*) = \{j \in I: \ \lambda^*_j > 0\};$$

(b) *it follows that*

$$d^T \nabla^2_{xx} L_{\rm mm}(x^*, \lambda^*) d > 0, \ \forall \ \lambda^* \in \Lambda_{\rm mm}(x^*), \ \forall \ d \in M(x^*, \lambda^*) \setminus \{0\};$$

(c) $\Lambda_{\rm mm}(x^*) = \{\lambda^*\}$ *is a singleton set and*

$$d^T \nabla^2_{xx} L_{\rm mm}(x^*, \lambda^*) d > 0, \ \forall \ d \in M(x^*, \lambda^*) \setminus \{0\},$$

then there exists a constant $\epsilon > 0$ such that

$$\overline{A}(x, \lambda) \equiv A(x^*), \ \text{i.e.,} \ J_{\rm mm}(x, \lambda) \equiv J(x^*), \ I_{\rm mm}(x, \lambda) \equiv I(x^*), \ \forall \ (x, \lambda) \in \mathcal{K}_{\rm mm}(x^*) + \mathcal{B}_{\epsilon}. \tag{3.6}$$

4. Numerical examples

To give the readers a feel for the potentialities of the proposed identifying technique, we use three minimax problems to compare our identification technique $\overline{A}(x,\lambda)$ to $A(x,\lambda)$ and $\hat{A}(x,\lambda)$ proposed in [18]. The test problems are given below, where "OF" means "objective functions" and "ICF" means "inequality constrained functions".

Test problem 01 [20, problem 6]. ICF: $f_4(x) = x_1^2 - x_2^2,\ f_5(x) = -2x_1^3 - x_2^3$; and

$$\text{OF}:\ f_1(x) = x_1^4 + x_2^2,\ f_2(x) = (2-x_1)^2 + (2-x_2)^2,\ f_3(x) = 2\exp(-x_1 + x_2).$$

The solution of this problem is $x^* = (1,1)^T$, and the LICQ is violated at x^*, but the MFCQ is satisfied at x^*. Again,

$$\Lambda_{\text{mm}}(x^*) = \left\{(t, 3t/2, 1-5t/2, 1-3t, 0)^T :\ t \in [0, 1/3]\right\}$$

is not a singleton. Furthermore, if $t \in \{0, 1/3\}$, then strict complementarity is also violated.

Test problems 02 [21] OF: $f_j(x) = x_j^2,\ j = 1,\ldots,10$; and

$$\text{ICF}:\ f_{10+j}(x) = (3 - 0.5x_{j+1})x_{j+1} - x_j - 2x_{j+2} + 1,\ j = 1,2,\ldots,8.$$

$$(x^*,\lambda^*) = (0.7071,\ldots,0.7071, 0.0279, 0.3208, 0, 0.6512, 0,0,0,0,0,0, 0.0395, 0.4605, 0,0,0,0,0,0).$$

Test problem 03 [21]. ICF: $f_{2+j}(x) = x_j^2 + x_{j+1}^2 + x_j x_{j+1} - 1,\ j = 1,2,\ldots,9$; and

$$\text{OF}:\ f_1(x) = \sum_{j=1}^{9}(x_j^2 + (x_{j+1}-1)^2 + x_{j+1} - 1),\ f_2(x) = \sum_{j=1}^{9}(-x_j^2 - (x_{j+1}-1)^2 + x_{j+1} + 1).$$
$$(x^*,\lambda^*) = (0,0,0,0,0,0,0,0,0,0, 0.75, 0.25, 0,0,0,0,0,0,0,0,0,0).$$

To test the three problems, random points (x,λ) at the neighborhood of its (approximately) known solution x^* with (approximate) $\lambda^* \in \Lambda_{\text{mm}}(x^*)$ were generated. More in detail, for each $\epsilon \in \{10^n, n = 1, 0, -1, -2\}$, we generated 500 random vectors (x,λ) on the boundary of the set

$$\mathcal{K}_{\text{mm}}(x^*) + \mathcal{B}_\epsilon = \{(x,\lambda) :\ \text{exists } \lambda^* \in \Lambda_{\text{mm}}(x^*) \text{ such that } \|(x,\lambda) - (x^*,\lambda^*)\|_\infty < \epsilon\}.$$

For each of these random vectors, we identified each constraint no matter it is active or inactive.

The computational results are reported in Tables 1–4 below. We compared our identification technique $\overline{A}(x,\lambda)$ to $A(x,\lambda)$ and $\hat{A}(x,\lambda)$ which were proposed in [18]. For each of the objective and constrained functions f_j as well as the different values of ϵ, we report the sum of the correctly identified objective functions and constraints over all 500 randomly generating vectors (x,λ), which is denoted by the notation nf_j. The last column of each table contains the total number of correctly identified constraints over all objective functions and constraints, which is denoted by "Total". The column of "Term" means the corresponding identification techniques.

The results reported in Table 4 further corroborate two important theory conclusions, i.e., the relations $\overline{A}(x,\lambda) \subseteq \hat{A}(x,\lambda) \subseteq A(x,\lambda)$ (these imply $|\overline{A}(x,\lambda)| \subseteq |\hat{A}(x,\lambda)| \subseteq |A(x,\lambda)|$) in all this cases, and $\overline{A}(x,\lambda) = \hat{A}(x,\lambda) = A(x,\lambda) \equiv A(x^*)$ if (x,λ) is sufficiently close to $\mathcal{K}_{\text{mm}}(x^*)$, e.g. for $\varepsilon = 0.01$. Thus, the proposed active set technique in this paper is a tighter accuracy identification for the active sets $J(x^*)$ and $I(x^*)$.

Table 1. Numerical results for test problems 01

Term	ϵ	nf_1	nf_2	nf_3	nf_4	nf_5	Total
$A(x,\lambda)$	10	345	0	157	461	372	1335
	1	396	0	168	359	491	1414
	0.1	500	431	500	500	500	2431
	0.01	500	500	500	500	500	2500
$\hat{A}(x,\lambda)$	10	325	0	175	478	352	1330
	1	398	0	184	363	485	1430
	0.1	500	439	500	500	500	2439
	0.01	500	500	500	500	500	2500
$\overline{A}(x,\lambda)$	10	304	78	159	331	263	1135
	1	308	218	216	500	292	1534
	0.1	461	448	486	500	500	2395
	0.01	500	500	500	500	500	2500

Table 2. Numerical results for test problem 02

Term	ϵ	nf_1	nf_2	nf_3	nf_4	nf_5	nf_6	nf_7	nf_8	nf_9	nf_{10}
$A(x,\lambda)$	10	67	80	88	83	72	67	70	66	73	72
	1	366	371	361	368	359	363	354	359	337	363
	0.1	500	500	500	500	500	500	500	500	500	500
	ϵ	nf_{11}	nf_{12}	nf_{13}	nf_{14}	nf_{15}	nf_{16}	nf_{17}	nf_{18}	Total	
	10	126	113	113	109	126	120	127	127	1699	
	1	414	424	427	417	438	418	428	419	6986	
	0.1	500	500	500	500	500	500	500	500	9000	
$\hat{A}(x,\lambda)$	ϵ	nf_1	nf_2	nf_3	nf_4	nf_5	nf_6	nf_7	nf_8	nf_9	nf_{10}
	10	62	64	75	69	72	90	78	72	72	73
	1	383	367	373	346	378	368	371	363	357	382
	0.1	500	500	500	500	500	500	500	500	500	500
	ϵ	nf_{11}	nf_{12}	nf_{13}	nf_{14}	nf_{15}	nf_{16}	nf_{17}	nf_{18}	Total	
	10	104	90	104	102	105	99	110	102	1543	
	1	318	351	294	332	320	331	314	307	6255	
	0.1	500	500	500	500	500	500	500	500	9000	
$\overline{A}(x,\lambda)$	ϵ	nf_1	nf_2	nf_3	nf_4	nf_5	nf_6	nf_7	nf_8	nf_9	nf_{10}
	10	72	67	73	69	76	76	77	70	91	71
	1	400	378	396	401	398	390	395	395	338	397
	0.1	500	500	500	500	500	500	500	500	500	500
	ϵ	nf_{11}	nf_{12}	nf_{13}	nf_{14}	nf_{15}	nf_{16}	nf_{17}	nf_{18}	Total	
	10	116	109	100	113	112	116	115	84	1607	
	1	326	317	296	326	318	311	316	371	6469	
	0.1	500	500	500	500	500	500	500	500	9000	

Table 3. Numerical results for test problems 03

Term	ϵ	nf_1	nf_2	nf_3	nf_4	nf_5	nf_6	nf_7	nf_8	nf_9	nf_{10}	nf_{11}	Total
$A(x,\lambda)$	10	500	0	0	0	0	0	0	0	0	0	0	500
	1	489	67	0	0	0	0	0	0	0	0	0	556
	0.1	438	471	393	394	391	395	389	393	392	392	391	4439
	0.01	500	500	500	500	500	500	500	500	500	500	500	5500
$\hat{A}(x,\lambda)$	10	500	0	413	429	413	410	415	416	419	414	423	4255
	1	486	77	20	14	18	16	18	18	18	19	14	718
	0.1	453	472	390	389	388	387	383	386	386	388	386	4408
	0.01	500	500	500	500	500	500	500	500	500	500	500	5500
$\overline{A}(x,\lambda)$	10	500	0	96	417	79	75	79	93	82	421	93	1935
	1	498	466	455	36	43	43	33	34	35	39	43	1725
	0.1	462	459	400	408	415	423	412	410	419	416	395	4619
	0.01	500	500	500	500	500	500	500	500	500	500	500	5500

Table 4. The sums of $|A(x,\lambda)|$, $|\hat{A}(x,\lambda)|$ and $|\overline{A}(x,\lambda)|$ for each ϵ and 500 random points

Term	P\\ϵ	10	1	0.1	0.01	P\\ϵ	10	1	0.1	0.01	P\\ϵ	10	1	0.1	0.01
$\lvert A(x,\lambda)\rvert$		1137	964	1933	2000		1659	6970	9000	9000		5000	5056	1875	1000
$\lvert \hat{A}(x,\lambda)\rvert$	01	1137	964	1933	2000	02	1655	6968	9000	9000	03	1279	1546	1875	1000
$\lvert \overline{A}(x,\lambda)\rvert$		1109	936	1895	2000		1598	6389	9000	9000		1259	1305	1781	1000

5. Conclusions

In this paper, we proposed a new identification technique for constrained nonlinear programming and minimax problems at an isolated local solution, which is tighter than the previous results and can be used to design effective penalty algorithms. As further work, we think there exist three problems are worthy of research. First, one can attempt to use this technique to design specific penalty algorithms. Second, one can try to define and construct new identification functions with better properties.

Third, noticing that the composition $\min\{F(x)-f_j(x), \lambda_j\}$ $(j \in J)$ of the function $\Phi_{mm}(x, \lambda)$ in (3.4) is a type of minimax function, its properties such as the Lipschitz continuity are not so good, we have attempted to structure a new function with better features to replace $\min\{F(x)-f_j(x), \lambda_j\}$. Let $a \geq 0$. Considering the following equivalent relations

$$\{b \geq 0,\ ab = 0\} \Longleftrightarrow \min\{0, b\} + ab = 0,$$

and $F(x) - f_j(x) \geq 0$ always holds, if one replaces $\min\{F(x)-f_j(x), \lambda_j\}$ $(j \in J)$ in the function $\Phi_{mm}(x, \lambda)$ defined in (3.4) by $\min\{0, \lambda_j\} + \lambda_j(F(x)-f_j(x))$ $(j \in J)$, whether the corresponding set $\overline{A}(x, \lambda)$ can accurately identify the active set $A(x^*)$? We think this is an interesting and **an open problem**.

References

[1] D.P. Bertsekas, Constrained optimization and Lagrange multiplier methods, Academic Press, New York, 1982.

[2] J. V. Burke, J. J. More, On the identification of active constraints, SIAM J. Numer. Anal. 25(1988) 1197–1211.

[3] J. Burke, On the identification of active constraints II: The nonconvex case, SIAM J. Numer. Anal. 27(1990), 1081–1102.

[4] M. Lescrenier, Convergence of trust region algorithms for optimization with bounds when strict complementarity does not hold, SIAM J. Numer. Anal. 28(1991) 476–495.

[5] S. J. Wright, Identifiable surfaces in constrained optimization, SIAM J. Control Optim. 31(1993) 1063–1079.

[6] W. Hare and A. Lewis, Identifying active constraints via partial smoothness and proxregularity, J. Convex Anal. 11(2004) 251–266

[7] A. S. Lewis, Active sets, nonsmoothness, and sensitivity, SIAM J. Optim. 13(2002) 702–725.

[8] R. D. C. Monteiro, S. J. Wright, Local convergence of interior-point algorithms for degenerate monotone LCP, Comput. Optim. Appl. 3(1994) 131–155.

[9] N. Yamashita, H. Dan, M. Fukushima, On the identification of degenerate indices in the nonlinear complementarity problem with the proximal point algorithm, Math. Program. Series A 99(2004) 377–397.

[10] N. H. Xiu, J. Z. Zhang, Identification of the optimal active set in a noninterior continuation method for LCP, J. Global Optim. 26(2003) 183–198.

[11] G. H. Lin, M. Fukushima, Hybrid approach with active set identification for mathematical programs with complementarity constraints, J. Optim. Theory Appl. 128(2006) 1–28.

[12] F. Facchinei, A. Fischer, C. Kanzow, On the accurate identification of active constraints, SIAM J. Optim. 9(1998) 14–32.

[13] C. Oberlin, S. J. Wright, Active set identification in nonlinear programming, SIAM J. Optim. 17(2006) 577–605.

[14] A. S. Lewis, S. J. Wright, Identifying activity, SIAM J. Optim. 21(2011) 597–614.

[15] J. B. Jian, Fast algorithms for smooth constrained optimization – theoretical analysis and numerical experiments, Science Press, Beijing, 2010.

[16] J. B. Jian, W. X. Cheng, A superlinearly convergent strongly sub-feasible SSLE-type algorithm with working set for nonlinearly constrained optimization, J. Comput. Appl. Math. 225(2009) 172–186.

[17] J. B. Jian, Y. Liu, A superlinearly convergent method of quasi-strongly sub-feasible directions with active set identifying for constrained optimization, Nonl. Anal.–RWA 12(2011) 2717–2729.

[18] D. L. Han, J. B. Jian, J. Li, On the accurate identification of active set for constrained minimax problems, Nonl. Anal.–TMA 74(2011) 3022–3032.

[19] P. Y. Nie, A new penalty method for nonlinear programming, Comput. Math. Appl. 52(2006) 883–896.

[20] B. Rustem, Q. Nguyen, An algorothm for the inequality-constrained discrete minimax problem, SIAM J. Optim. 8(1998) 265–283.

[21] Napsu Karmitsa, Test problems for large-scale nonsmooth minimization, Reports of the Department of Mathematical Information Technology, Series B, Scienctific Computing, No. B.4, 2007.