A Proximal Method for Identifying Active Manifolds

W.L. Hare*

April 18, 2006

Abstract

The minimization of an objective function over a constraint set can often be simplified if the "active manifold" of the constraints set can be correctly identified. In this work we present a simple subproblem, which can be used inside of any (convergent) optimization algorithm, that will identify the active manifold of a "prox-regular partly smooth" constraint set in a finite number of iterations.

Key words: Nonconvex Optimization, Active Constraint Identification, Prox-regular, Partly Smooth
AMS 2000 Subject Classification:
Primary: 90C26, 65K05
Secondary: 26B25, 49M99, 52A30

1 Introduction

The ideas and inspiration behind the study of active manifold identification can largely be traced back to the following problem:

$$\min\{f(x) : g_i(x) \leq 0,\ i = 1, 2, ...N\} \tag{1}$$

where each function g_i is twice continuously differentiable. Clearly it would be of great advantage to optimizers to know exactly which of the functions g_i were "active" at the minima of the problem. That is, if for a minima of the problem $\bar{x}$ one had access to the *active set*

$$I := \{i : g_i(\bar{x}) = 0\},$$

one could simplify the problem by focusing on optimizing over the corresponding *active manifold*

$$\mathcal{M}_I := \{x : g_i(x) = 0, i \in I,\ g_i(x) < 0, i \notin I\}.$$

*Department of Computing and Software, McMaster University, Hamilton, ON L8N 3Z5, Canada, `whare@cecm.sfu.ca`

Before motivating further, let us recall that for a convex set C the *convex normal cone* to C at the point $\bar{x} \in C$ is

$$\{n : \langle n, x - \bar{x}\rangle \leq 0 \text{ for all } x \in C\}.$$

Using the normal cone notation, the first order optimality conditions for the minimization problem $\min_x\{f(x) : x \in C\}$ can be written $-\nabla f(\bar{x}) \in N_C(\bar{x})$. As such, points where this holds are refereed to as *critical points*. We call $\bar{x}$ a *nondegenerate critical point* to the problem if $-\nabla f(\bar{x}) \in \text{rint}\, N_C(\bar{x})$, where rint represents the relative interior of a set (the interior relative to the smallest affine set containing the set in question). A *strict critical point* is a point $\bar{x}$ with $-\nabla f(\bar{x}) \in \text{int} N_C(\bar{x})$, where int represents the interior of a set.

It is well known that, if the constraint set C and the objective function f are convex, then any strict critical point is the unique minimizer to f over C (see [12, Thm 6.12] for example). Furthermore, one can easily show that, if C is defined by a finite number of smooth constraints, $C := \{x : g_i(x) \leq 0,\ i = 1, 2, ...N\}$ then the active manifold for C is the singleton $\{\bar{x}\}$ (see Example 2.3 below); so identifying the active manifold is equivalent to solving the minimization problem. In 1991, Al-Khayyal and Kyparisis proposed the following rather elegant method for identifying the active manifold in this case:

Theorem 1.1 *[1, Thm 2.1] Suppose the point $\bar{x}$ is a strict critical point to the minimization problem $\min_x\{f(x) : x \in C\}$ where $C := \{x : g_i(x) \leq 0,\ i = 1, 2, ...N\}$ and the functions f and g_i all convex continuously differentiable. If x_k converges to $\bar{x}$, then for all k sufficiently large*

$$\text{argmin}_p\{\langle \nabla f(x_k), p\rangle\} = \{\bar{x}\}. \tag{2}$$

The proof is straight forward and can be argued as follows (for details see [1]):

> Since $\nabla f(x_k) \to \nabla f(\bar{x})$ (as $x_k \to \bar{x}$ and ∇f is continuous) and $-\nabla f(\bar{x}) \in \text{int} N_C(\bar{x})$, eventually $-\nabla f(x_k) \in \text{int} N_C(\bar{x})$. At this time $\bar{x}$ becomes a strict critical point for Subproblem (2), and hence its unique minimizer. □

In reviewing the proof of [1, Thm 2.1] it is immediately clear that many of the theorem's conditions can be relaxed. Al-Khayyal and Kyparisis do exactly that in [1, Thm 3.1], replacing $\nabla f(x_k)$ with an arbitrary convergent sequence and generalizing to any convex constraint set. However, Al-Khayyal and Kyparisis were unable to remove the requirement that $\bar{x}$ is a strict critical point to the problem, nor the requirement that the constraint set is convex.

In this paper we alter Subproblem (2) by adding a quadratic barrier function and show how this allows active manifold identification on a much broader collection of constraint sets. In particular, we successfully replace the strict critical point assumption with the weaker assumption that $\bar{x}$ is a nondegenerate critical point, remove all conditions on the objective function, and reduce the restrictions on the constraint set from convex to *prox-regular* (see [10]; Definition 1.2 in this work). To do this we make use of the framework of *partly smooth sets* (see [7]; Definition 2.1 in this work), and the recent manifold identification results found in [4].

The remainder of this paper is organized as follows. In the next subsection we outline our notation, as well as provide the required definitions and background for this work. In Section 2 we formally define the framework for the active manifolds which we use in this paper, *partly smooth sets*, and show that a small alternation to Subproblem (2) allows us to identify the active manifold of partly smooth sets. In particular, Theorem 2.7 shows how to achieve active manifold identification in a general framework, while Example 2.8 returns this result to an optimization setting. We conclude, in Section 3, with two examples showing that the active manifold identification results cannot be achieved without the rewriting of Subproblem (2).

1.1 Notation and Definitions

In general we shall follow the notation laid out by Rockafellar and Wets in [12].

In particular, a vector n is considered *normal* to a set S at a point $\bar{x} \in S$ *in the regular sense* if

$$\liminf_{x \to \bar{x},\ x \in S \setminus \bar{x}} \frac{\langle n, x - \bar{x} \rangle}{|x - \bar{x}|} \leq 0,$$

we denote the cone of all such vectors by $\hat{N}_S(\bar{x})$. The *limiting normal cone* (a.k.a. the Clarke normal cone) to S at $\bar{x}$ is the collection

$$N_S(\bar{x}) := \limsup_{x \to \bar{x}, x \in S} \hat{N}_S(x).$$

We say the set S is regular at $\bar{x}$ if $\hat{N}_S(\bar{x}) = N_S(\bar{x})$ [12, Def 6.3 & 6.4]. If the set S is convex, then S is regular at all points $x \in S$, and the regular normal cone reduces to the convex normal cone described in Section 1 of this work [12, Thm 6.9]. *Critical points*, *nondegenerate critical points*, and *strict critical points* are all defined for nonconvex sets in terms of the limiting normal cone in the obvious manner.

The concept of a normal vector is closely related to the *projection* of a point $\bar{x}$ onto a closed set S: $\mathrm{P}_S(\bar{x}) := \mathrm{argmin}\{|y - \bar{x}| : y \in S\}$. A vector $\bar{n}$ is a *proximal normal vector* to S at $\bar{x}$ if for some $\bar{r} > 0$ the projection of $\bar{x} + \frac{1}{\bar{r}}\bar{n}$ onto S is equal to $\bar{x}$. In this case, the projection of $\bar{x} + \frac{1}{r}\bar{n}$ onto S will be $\bar{x}$ for any $r > \bar{r}$, and the infimum of all such $\bar{r}$ is called the *projection threshold* for $\bar{n}$ [12, Ex 6.16].

It is clear that any normal vector to a convex set is a proximal normal vector with a projection threshold of 0. A similar result holds for the much broader class of sets deemed *prox-regular*. Although prox-regularity was first introduced in terms of functions [10] [9] and generalized to sets via indicator functions, in [11] it was shown that for sets the following definition suffices.

Definition 1.2 (Prox-regular sets) *A closed set $S \subseteq \mathbf{R}^n$ is* prox-regular *at a point $\bar{x} \in S$ if the projection mapping is single valued near $\bar{x}$.*

The definition of prox-regular sets makes it clear that all convex sets are prox-regular. Moreover, in [11, Cor 2.2] it was shown that if the set S is prox-regular at the point $\bar{x}$, then any normal vector to S at $\bar{x}$ is a proximal normal vector.

2 Identifying Partly Smooth Manifolds

Recent years have seen a good deal of research which extends the idea of the active manifold for finitely constrained sets to a broader more manageable class. For example, [2] explores the idea of *open facets*, a generalization of polyhedral faces to any surface of a set that locally appears flat; while [13] develops the idea of $\mathcal{C}^p$*-identifiable surfaces*, surfaces of sets which can be nicely described via a finite number of constraint even when the set cannot. In [7] the idea of a *partly smooth function* is developed, and from it the notion of a *partly smooth set* (see Definition 2.1 below). In this work we chose to focus on partly smooth sets for three main reasons.

First, in [3] it was shown that the class of partly smooth sets contains many of the recently developed classes of sets containing active manifolds (such as open facets and identifiable surfaces). Second, unlike many of the most other classes, partial smoothness does not invoke convexity in its definition. As such, partial smoothness provides a very broad framework for our results.

Our third reason lies in the recent results of [4], which describes exactly what is required to identify the active manifold of a partly smooth set [4, Thm 4.1]. For the reader's convenience we restate this result in Theorem 2.4 below.

Next we formally define partly smooth sets.

Definition 2.1 (Partly smooth) *A set* $S \subset \mathbf{R}^m$ *is* partly smooth *at a point* $\bar{x} \in S$ *relative to a set* $\mathcal{M} \subseteq S$ *if* $\mathcal{M}$ *is a smooth* ($\mathcal{C}^2$) *manifold about* $\bar{x}$ *and the following properties hold:*

(i) $S \cap \mathcal{M}$ *is a neighbourhood of* $\bar{x}$ *in* $\mathcal{M}$*;*

(ii) S *is regular at all points in* $\mathcal{M}$ *near* $\bar{x}$*;*

(iii) $N_{\mathcal{M}}(\bar{x}) \subseteq N_S(\bar{x}) - N_S(\bar{x})$*; and*

(iv) *the normal cone map* $N_S(\cdot)$ *is continuous at* $\bar{x}$ *relative to* $\mathcal{M}$*.*

We then refer to $\mathcal{M}$ *as the* active manifold *(of partial smoothness).*

Before examining our method of identifying the active manifold for a partly smooth constraint set, we provide two simple examples which draw connections between prox-regular partly smooth sets and constrained optimization. Our first example examines sets formed via a finite number of smooth constraints.

Example 2.2 (Finitely constrained sets) Consider the set

$$S := \{x : g_i(x) \leq 0,\ i = 1, 2, \ldots, n\},$$

where $g_i \in \mathcal{C}^2$.

For any point $\bar{x} \in S$ define $I(\bar{x}) := \{i : g_i(\bar{x}) = 0\}$. If the *active gradients* of S at $\bar{x}$, $\{\nabla g_i(\bar{x}) : i \in I(\bar{x})\}$, form a linearly independent set, then S is prox-regular at $\bar{x}$ and partly smooth there relative to the active manifold

$$\mathcal{M}_g := \{x : g_i(x) = 0, i \in I,\ g_i(x) < 0, i \notin I\}$$

([10, Cor 2.12] and [7, 6.3]). □

A second example of prox-regular partial smoothness is generated by examining strict critical points.

Example 2.3 (Strict critical points) If the set $S \subseteq \mathbf{R}^n$ is regular at the point $\bar{x} \in S$ and the normal cone $N_S(\bar{x})$ has interior, then S is partly smooth at $\bar{x}$ relative to the manifold $\{\bar{x}\}$.

Indeed, as $\{\bar{x}\}$ is a singleton conditions (i) and (iv) hold true. Condition (ii) is given, while condition (iii) follows from $N_{\mathcal{M}}(\bar{x}) = \mathbf{R}^n$ and $N_S(\bar{x})$ having interior. □

The primary goal of this work is examine a new method to identify the active manifold of a partly smooth constraint set. In order to maintain a general setting, we shall consider the following situation. Consider:

a constraint set S,

a sequence of points $x_k \in S$ which converge to the point $\bar{x} \in S$, and

a sequence of vectors d_k which converge to the normal vector $\bar{n} \in -N_S(\bar{x})$.

In an optimization sense, the points x_k might represent a sequence of iterates generated from an optimization algorithm. In this case, the point $\bar{x}$ would be the minima (or maxima) of the function, the vectors d_k would be a sequence of gradient or subgradient vectors, while $\bar{n}$ would represent the first order optimality conditions for the problem. Notice that, in our general framework, we make no assumptions about how these iterates are generated, nor about the objective function of the optimization problem.

Using the points x_k and vectors d_k we create the following subproblem,

$$p_k \in \operatorname{argmin}_p\{\langle d_k, p\rangle + r\frac{1}{2}|p - x_k|^2 : p \in S\}. \tag{3}$$

In the main result of this paper we show that, not only does p_k converge to $\bar{x}$, but p_k identifies the active manifold of the constraint set in a finite number of iterations.

Assuming that the vectors d_k represent gradient or subgradient vectors, Subproblem (3) can be thought of in several manners. In some sense, p_k represents a proximal point for the linear function $\langle d_k, \cdot\rangle$. As such, p_k could be seen as the result of taking exactly one "null step" of a bundle method applied to the objective function [6]. In another sense, by rewriting $r\frac{1}{2}|p - x_k|^2$ as $\langle p - x_k, rI(p - x_k)\rangle$, the points p_k could be seen as quasi-Newton steps using the approximate Hessian matrix $H_k \equiv rI$ for all k [8, Chpt 6 & 8]. We prefer to think of p_k in the first sense, as in quasi-Newton methods one expects the approximate Hessian to somehow represent the correct Hessian of the objective function at the point x_k.

In order to show Subproblem (3) correctly identifies the active manifold, we shall make use of the following result by Hare and Lewis.

Theorem 2.4 *Consider a set S that is partly smooth at the point $\bar{x}$ relative to the manifold $\mathcal{M}$ and prox-regular at $\bar{x}$. If the vector $\bar{n}$ satisfies $-\bar{n} \in \operatorname{rint} N_S(\bar{x})$ and the sequences $\{x_k\}$ and $\{d_k\}$ converge to $\bar{x}$ and $\bar{n}$ respectively, then*

$$\operatorname{dist}(-d_k, N_S(x_k)) \to 0 \quad \Leftrightarrow \quad x_k \in \mathcal{M} \;\; \textit{for all large } k.$$

Proof: ($\Rightarrow$) See [4, Thm 4.1].

($\Leftarrow$) (From [5, Thm 2.1]) Note if $x_k \in \mathcal{M}$ for all k large, then condition (iv) of partial smoothness implies $N_S(x_k) \to N_S(\bar{x})$. Applying regularity (condition (ii) of partial smoothness) and [12, Cor 4.7] we see

$$\operatorname{dist}(d_k, N_S(x_k)) = \operatorname{dist}(0, N_S(x_k) - d_k) \to \operatorname{dist}(0, N_S(\bar{x}) - \bar{n}) = 0.$$

□

In order to apply [4, Thm 4.1] we first require a lemma which bounds the points Subproblem (3) creates.

Lemma 2.5 (Bounding p_k) *Consider a closed set S, a point $x \in S$, direction d and a parameter $r > 0$. If $\bar{p}$ is defined via Subproblem (3) (with x_k, d_k and p_k replaced with x, d, and $\bar{p}$), then*

$$|\bar{p} - x + d/(2r)| \leq \frac{|d|}{2r}.$$

Proof: For a given point x and direction d, consider the problem

$$\begin{aligned}
&\operatorname{argmin}_p\{\langle d, p\rangle + r\tfrac{1}{2}|p - x|^2 : p \in S\} \\
&= \operatorname{argmin}_{p=x+\alpha w}\{\langle d, x + \alpha w\rangle + r\tfrac{1}{2}|\alpha w|^2 : |w| = 1,\ \alpha \geq 0, p \in S\} \\
&= \operatorname{argmin}_{p=x+\alpha w}\{\langle d, \alpha w\rangle + r\tfrac{1}{2}\alpha^2 : |w| = 1,\ \alpha \geq 0, p \in S\}.
\end{aligned}$$

For a given $w \in \mathbf{R}^n$ with $|w| = 1$, the function

$$\langle d, w\rangle\alpha + r\frac{1}{2}\alpha^2$$

is a one dimensional quadratic function which is decreasing from $\alpha = 0$ to $\alpha = \langle -d, w\rangle/r$. Therefore

$$\begin{aligned}
\operatorname{argmin}_p\{\langle d, p\rangle + r\tfrac{1}{2}|p - x|^2 : p \in S\} &\subseteq \{x + \alpha w : |w| = 1,\ 0 \leq \alpha \leq \langle -d, w\rangle/r\} \\
&= \{x + y : |y| = \alpha,\ 0 \leq \alpha^2 \leq \langle -d, y\rangle/r\} \\
&= \{x + y : |y|^2 \leq \langle -d, y\rangle/r\} \\
&= \{x + y : |y|^2 + \langle d/r, y\rangle + |\tfrac{d}{2r}|^2 \leq |\tfrac{d}{2r}|^2\} \\
&= \{x + y : |y + d/(2r)|^2 \leq |\tfrac{d}{2r}|^2\} \\
&= \{p : |p - x + d/(2r)| \leq \tfrac{|d|}{2r}\},
\end{aligned}$$

which proves the lemma. □

Now that we have a bound for the points p_k created by Subproblem (3), we can show that all feasible points within this bound must converge to $\bar{x}$.

Lemma 2.6 *Let $-\bar{n}$ be a proximal normal to the closed set S at the point $\bar{x}$. Consider any two sequences $x_k \to \bar{x}$ and $d_k \to \bar{n}$, a parameter $r > 0$. If r is sufficiently large that $\mathrm{P}_S(\bar{x} - \bar{n}/(2r)) = \{\bar{x}\}$, then*

$$\max_p\{|\bar{x} - p| : |x_k - d_k/(2r) - p| \leq |d_k|/2r,\ p \in S\} \to 0.$$

Proof: We begin by noting that the max is well defined as for each k the set

$$S_k := \{p : |x_k - d_k/(2r) - p| \leq |d_k|/2r,\ p \in S\}$$

is closed and bounded. We may therefore for each k find some

$$p_k \in \text{argmax}_p\{|\bar{x} - p| : |x_k - d_k/(2r) - p| \leq |d_k|/2r,\ p \in S\}.$$

Note that p_k forms a bounded sequence as

$$\begin{aligned} |\bar{x} - p_k| &\leq |\bar{x} - x_k| + |x_k - d_k/(2r) - p| + |d_k/(2r)| \\ &\leq |\bar{x} - x_k| + 2|d_k/(2r)|, \end{aligned}$$

and $|\bar{x} - x_k| + 2|d_k/(2r)|$ converges to $|\bar{n}|/r$.

Dropping to a subsequence as necessary we assume that p_k converges to a point $\bar{p}$. By definition we know that $|(x_k - d_k/(2r) - p_k| \leq |d_k|/(2r)$ and $p_k \in S$ for all k. Passing to the limit, noting S is closed, we find

$$|\bar{x} - \bar{n}/(2r) - \bar{p}| \leq |\bar{n}|/(2r) \text{ and } \bar{p} \in S.$$

By our assumptions on $\bar{n}$ we know that

$$\begin{aligned} \min\{|\bar{x} - \bar{n}/(2r) - y| : y \in S\} &= |\bar{n}|/(2r) \\ \text{argmin}\,\{|\bar{x} - \bar{n}/(2r) - y| : y \in S\} &= \{\bar{x}\}. \end{aligned}$$

Therefore $\bar{x} = \bar{p}$, and we conclude that

$$\max_p\{|\bar{x} - p| : |x_k - d_k/(2r) - p| \leq |d_k|/2r,\ p \in S\} \to 0.$$

□

Before continuing it is worth remarking on the use of proximal normals in Lemma 2.6. In order to ensure the existence of a parameter r sufficiently large to control the projection of $\bar{x} - \bar{n}/(2r)$ onto the constraint set, Lemma 2.6 assumes $-\bar{n}$ is a proximal normal to the constraint set S at $\bar{x}$. In the case of S being a prox-regular set, all normals are proximal normals, so we no longer need to state "$-\bar{n}$ is a proximal normal to the constraint set S at the point $\bar{x}$." If the constraint set S is convex then all normals are proximal normals with threshold 0, so "r sufficiently large" reduces to "$r > 0$".

We now turn to the major result in this work, which states that Subproblem (3) identifies the active manifold.

Theorem 2.7 (Identifying the active manifold) *Consider a constraint set S that is partly smooth at the point $\bar{x}$ relative to the manifold $\mathcal{M}$ and prox-regular at $\bar{x}$. Suppose the vector $\bar{n}$ satisfies $-\bar{n} \in \text{rint}\, N_S(\bar{x})$, and the sequences $\{x_k\}$ and $\{d_k\}$ converge to $\bar{x}$ and $\bar{n}$ respectively. Fix a parameter $r > 0$ and define a sequence*

$$p_k \in \text{argmin}_p\{\langle d_k, p\rangle + r\frac{1}{2}|p - x_k|^2 : p \in S\}.$$

If r sufficiently large then

$$p_k \in \mathcal{M} \quad \textit{for all } k \textit{ large.}$$

Proof: First note that, by prox-regularity, the vector $-\bar{n}$ is a proximal normal vector to S at $\bar{x}$; therefore we may assume r is sufficiently large for $\mathrm{P}_S\left(\bar{x} - \bar{n}/(2r)\right) = \{\bar{x}\}$ to hold.

We begin the proof by showing that p_k converges to $\bar{x}$. Indeed by Lemma 2.5 we have for all k that

$$p_k \in \{p : |x_k - d_k/(2r) - p| \leq |d_k|/2r,\ p \in S\}.$$

By Lemma 2.6 we know that

$$\max\{|\bar{x} - p| : |x_k - d_k/(2r) - p| \leq |d_k|/2r,\ p \in S\} \to 0.$$

Thus we must have $|\bar{x} - p_k| \to 0$.

For each k define the vector $n_k := d_k + r(p_k - x_k)$. Passing to a limit on n_k we see

$$-n_k \to -\bar{n} - r(\bar{x} - \bar{x}) = -\bar{n}.$$

By the optimality of p_k we know that $-n_k \in N_S(p_k)$ for each k. Therefore we have

$$\mathrm{dist}(-n_k, N_S(p_k)) \equiv 0.$$

Applying Theorem 2.4 completes the proof. □

It is now an easy exercise to see how our result on active manifold identification relates to a specific instance of optimization.

Example 2.8 (Identifying active constraints) Consider a minimization problem,

$$\min\{f(x) : g_i(x) \leq 0, i = 1, 2, ...N\}, \tag{4}$$

where $f \in \mathcal{C}^1$ and $g_i \in \mathcal{C}^2$ for $i = 1, 2, ...N$. Define $S := \{x : g_i(x) \leq 0, i = 1, 2, ...N\}$, and suppose at the point $\bar{x} \in S$. If the active gradients,

$$\{\nabla g_i(\bar{x})\}_{i \in I} \quad \text{where } I = \{i : g(\bar{x}) = 0\}$$

form a linearly independent set, then Example 2.2 shows that S is prox-regular at $\bar{x}$ and partly smooth there relative to the active manifold

$$\mathcal{M} := \{x : g_i(x) = 0 \text{ for } i \in I,\ \text{ and } g_i(x) < 0 \text{ for } i \notin I\}.$$

Therefore, if $\bar{x}$ is a nondegenerate critical point of (4) (i.e. $-\nabla f(\bar{x}) \in \mathrm{rint}\, N_S(\bar{x})$) and the sequence x_k converges to $\bar{x}$, then for r sufficiently large

$$\mathrm{argmin}_p\{\langle \nabla f(x_k), p\rangle + r\frac{1}{2}|p - x_k|^2 : p \in S\} \in \mathcal{M} \quad \text{for all } k \text{ large.}$$

In Corollary 2.9 we shall see, if the set S is convex, then "r sufficiently large" reduces to "$r > 0$". □

We conclude this section with two corollaries showing how the conditions of Theorem 2.7 can be simplified if the constraint set S is convex (Corollary 2.9) or if the vector $\bar{n}$ is a strict critical point (Corollary 2.10).

Corollary 2.9 (Convex case) *Consider a convex set C that is partly smooth at the point $\bar{x}$ relative to the manifold $\mathcal{M}$. Suppose the vector $\bar{n}$ satisfies $-\bar{n} \in \operatorname{rint} N_C(\bar{x})$, and the sequences $\{x_k\}$ and $\{d_k\}$ converge to $\bar{x}$ and $\bar{n}$ respectively. Fix a parameter $r > 0$ and define a sequence p_k via Subproblem (3). Then*

$$p_k \in \mathcal{M} \quad \textit{for all } k \textit{ large.}$$

Proof: Convexity shows $\bar{n}$ is a proximal normal vector with threshold 0, and also provides the prox-regularity needed to apply Theorem 2.4. □

Corollary 2.10 (Strict critical point case) *Consider a set S which is prox-regular at the point $\bar{x}$. Suppose the vector $\bar{n}$ satisfies $-\bar{n} \in \operatorname{int} N_S(\bar{x})$, and the sequences $\{x_k\}$ and $\{d_k\}$ converge to $\bar{x}$ and $\bar{n}$ respectively. Fix a parameter $r > 0$ and define a sequence p_k via Subproblem (3). If r sufficiently large then*

$$p_k = \bar{x} \quad \textit{for all } k \textit{ large.}$$

Proof: Example 2.3 shows that the set is partly smooth at $\bar{x}$ relative to the manifold $\{\bar{x}\}$, and Theorem 2.7 completes the proof. □

3 Two Examples

As mentioned, a large inspiration for this work was the earlier results of [1]. Theorem 1.1 of this work restates and gives a brief proof of their main result.

It is worth noting that, it is possible to reestablish the results of [1] via the Lemmas and Theorems in this work. In doing this, one would have to rewrite Lemma 2.5 with $r = 0$, which results in bounding p_k in a half space instead of a ball. Lemma 2.6 can then be rewritten for $r = 0$ by making strong use of $\bar{x}$ being a strict critical point. Finally, Theorem 2.7 would have to be rewritten for $r = 0$, which would yield [1, Thm 3.1]. However, given the elegance of the original proof, we shall not go through the details here. Instead we shall show that without the strict critical point assumption, this active manifold identification technique cannot work for $r = 0$.

We do this via two examples. The first shows that, when $\bar{x}$ is not a strict critical point, the identified constraints may include constraints that are inactive at $\bar{x}$. In the second we show that, again when $\bar{x}$ is not a strict critical point, the identified constraints may fail to include the actual active constraints at $\bar{x}$.

Our first example also shows that without $r > 0$, Subproblem (3) can actually move iterates off of the correct active manifold.

Example 3.1 (False positive identification) Consider the convex box set

$$\begin{aligned} C &= \{(x,y) : -1 \leq x \leq 1, -1 \leq y \leq 1\}, \\ &= \{(x,y) : g_i(x,y) \leq 0, i = 1,2,3,4\}, \end{aligned}$$

where $g_1(x,y) = -x-1$, $g_2(x,y) = x-1$, $g_3(x,y) = -y-1$ and $g_4(x,y) = y-1$. Then C is partly smooth at the point $\bar{x} = (1,0)$ relative to the manifold $\mathcal{M} = \{(x,y) : g_2(x,y) = 0, \quad g_i(x,y) < 0 \ i = 1,3,4\}$. That is, the only active constraint at the point $\bar{x}$ is g_2.

Consider now the nondegenerate normal vector $-\bar{n} = (1,0) \in N_C(\bar{x}), \ (\bar{n} = (-1,0))$, along with the sequence of points $x_k = (1, 1/k)$ which converge to $\bar{x}$ and the sequence of direction vectors $d_k = (-1, 1/k)$ which converge to $\bar{n}$. These points and direction vectors satisfy the conditions of Corollary 2.9, so the solution to

$$p_k \in \operatorname{argmin}\{\langle d_k, p\rangle + r/2|p - x_k|^2 : p \in C\},$$

eventually lies on $\mathcal{M}$ for any $r > 0$. Indeed, one finds $p_k = (1, \frac{r-1}{rk}) \in \mathcal{M}$ for all k. However, the problem

$$p_k \in \operatorname{argmin}\{\langle d_k, p\rangle : p \in C\},$$

is solved at $p_k = (1, -1) \notin \mathcal{M}$ for all k. At this point both g_2 and g_3 are active. □

Example 3.2 (False negative identification) Consider the convex set

$$\begin{aligned} C &= \{(x,y) : -1 \leq x \leq 1, y \leq {}^{+}\sqrt{1-x^2}\}, \\ &= \{(x,y) : g_i(x,y) \leq 0, i = 1,2,3,4\}, \end{aligned}$$

where $g_1(x,y) = -x-1$, $g_2(x,y) = x-1$ and $g_3(x,y) = y - {}^{+}\sqrt{1-x^2}$. Then C is partly smooth at the point $\bar{x} = (1,-1)$ relative to the manifold $\mathcal{M} = \{(x,y) : g_2(x,y) = 0, \quad g_i(x,y) < 0 \ i = 1,3\}$. That is, the only active constraint at the point $\bar{x}$ is g_2.

Consider now the nondegenerate normal vector $-\bar{n} = (1,0) \in N_C(\bar{x}), \ (\bar{n} = (-1,0))$, along with the sequence of points $x_k = (1, -1 + 1/k)$ which converge to $\bar{x}$ and the sequence of direction vectors $d_k = (-1, -1/k)$ which converge to $\bar{n}$. These points and direction vectors satisfy the conditions of Corollary 2.9, so the solution to

$$p_k \in \operatorname{argmin}\{\langle d_k, p\rangle + r/2|p - x_k|^2 : p \in C\},$$

eventually lies on $\mathcal{M}$ for any $r > 0$. Indeed, one finds $p_k = (1, \frac{r+1}{rk} - 1)$ for all $k > \frac{r+1}{r}$. However the problem

$$p_k \in \operatorname{argmin}\{\langle d_k, p\rangle : p \in C\},$$

is solved at $p_k = (\frac{1}{\sqrt{1+k^{-2}}}, \frac{1}{k\sqrt{1+k^{-2}}}) \notin \mathcal{M}$ for all k. At this point g_3 is active instead of g_2. □

References

[1] F. Al-Khayyal and J. Kyparisis. Finite convergence of algorithms for nonlinear programs and variational inequalities. *J. Optim. Theory Appl.*, 70(2):319–332, 1991.

[2] J. V. Burke and J. J. Moré. On the identification of active constraints. *SIAM J. Numer. Anal.*, 25(5):1197–1211, 1988.

[3] W. L. Hare. Functions and sets of smooth substructure: relationships and examples. *Comput. Optim. Appl.*, 33(2-3):249–270, 2006.

[4] W. L. Hare and A. S. Lewis. Identifying active constraints via partial smoothness and prox-regularity. *J. Convex Anal.*, 11(2):251–266, 2004.

[5] W. L. Hare and A. S. Lewis. Identifying active manifolds. *submitted to: J. Optim. Theory Appl.*, pages 1–14, 2006. preprint available at: `www.cecm.sfu.ca/∼whare`.

[6] J.B. Hiriart-Urruty and C. Lemaréchal. *Convex Analysis and Minimization Algorithms. II*, volume 306 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1993. Advanced theory and bundle methods.

[7] A. S. Lewis. Active sets, nonsmoothness, and sensitivity. *SIAM J. Optim.*, 13(3):702–725 (electronic) (2003), 2002.

[8] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer Series in Operations Research. Springer-Verlag, New York, 1999.

[9] R. A. Poliquin and R. T. Rockafellar. Generalized hessian properties of regularized nonsmooth functions. *SIAM J. Optim.*, 6(4):1121–1137, 1996.

[10] R. A. Poliquin and R. T. Rockafellar. Prox-regular functions in variational analysis. *Trans. Amer. Math. Soc.*, 348(5):1805–1838, 1996.

[11] R. A. Poliquin, R. T. Rockafellar, and L. Thibault. Local differentiability of distance functions. *Trans. Amer. Math. Soc.*, 352(11):5231–5249, 2000.

[12] R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1998.

[13] S. J. Wright. Identifiable surfaces in constrained optimization. *SIAM J. Control Optim.*, 31(4):1063–1079, 1993.