Computational Optimization and Applications, 17, 279–299, 2000

On Two Applications of *H*-Differentiability to Optimization and Complementarity Problems

M.A. TAWHID tawhid@math.umbc.edu
M. SEETHARAMA GOWDA gowda@math.umbc.edu
Department of Mathematics and Statistics, University of Maryland, Baltimore County, Baltimore, MD 21250, USA

Received February 12, 1999; Accepted October 27, 1999

Abstract. In a recent paper, Gowda and Ravindran (Algebraic univalence theorems for nonsmooth functions, Research Report, Department of Mathematics and Statistics, University of Maryland, Baltimore, MD 21250, March 15, 1998) introduced the concepts of H-differentiability and H-differential for a function $f : R^n \to R^n$ and showed that the Fréchet derivative of a Fréchet differentiable function, the Clarke generalized Jacobian of a locally Lipschitzian function, the Bouligand subdifferential of a semismooth function, and the C-differential of a C-differentiable function are particular instances of H-differentials.

In this paper, we consider two applications of H-differentiability. In the first application, we derive a necessary optimality condition for a local minimum of an H-differentiable function. In the second application, we consider a nonlinear complementarity problem corresponding to an H-differentiable function f and show how, under appropriate conditions on an H-differential of f, minimizing a merit function corresponding to f leads to a solution of the nonlinear complementarity problem. These two applications were motivated by numerous studies carried out for C^1, convex, locally Lipschitzian, and semismooth function by various researchers.

Keywords: H-differentiability, nonlinear complementarity problem, NCP function, merit function, locally Lipschitzian function, generalized Jacobian

1. Introduction

In a recent paper [10], Gowda and Ravindran introduced the concepts of the H-differentiability and H-differential for a function $f : R^n \to R^n$. They showed that Fréchet differentiable (locally Lipschitzian, semismooth, C-differentiable) functions are H-differentiable (at given $\bar{x}$) with an H-differential given by $\{\nabla f(\bar{x})\}$ (respectively, $\partial f(\bar{x})$, $\partial_B f(\bar{x})$, C-differential). In their paper, Gowda and Ravindran investigated the injectivity of an H-differentiable function based on conditions on H-differentials. Also, in [25], H-differentials were used to characterize $\mathbf{P}(\mathbf{P}_0)$-functions.

In this paper, we consider two applications of H-differentiability. In the first application, we derive a necessary optimality condition for a local minimum of an H-differentiable real valued function. Specifically, we show in Theorem 3 that if x^* is a local minimum of such a function f, then $0 \in \overline{co}\, T_f(x^*)$ where $T_f(x^*)$ denotes an H-differential of f at x^*.

In the second application, we consider a nonlinear complementarity problem NCP(f) corresponding to an H-differentiable function $f : R^n \to R^n$: Find $\bar{x} \in R^n$ such that

$$\bar{x} \geq 0, \quad f(\bar{x}) \geq 0 \quad \text{and} \quad \langle f(\bar{x}), \bar{x} \rangle = 0.$$

By considering an NCP function $\Phi : R^n \to R^n$ associated with NCP(f) so that

$$\Phi(\bar{x}) = 0 \Leftrightarrow \bar{x} \text{ solves NCP}(f),$$

and the corresponding merit function

$$\Psi(x) := \frac{1}{2}\|\Phi(x)\|^2, \tag{1}$$

in this paper (see Sections 6, 7, and 8), we show how, under appropriate $\mathbf{P}_0$($\mathbf{P}$, regularity)-conditions on an H-differential of f, finding local/global minimum of Ψ (or a 'stationary point' of Ψ) leads to a solution of the given nonlinear complementarity problem. Our results unify/extend various similar results proved in the literature for C^1, locally Lipschitzian, and semismooth functions [1, 5–9, 11–14].

2. Preliminaries

We regard vectors in R^n as column vectors. We denote the inner-product between two vectors x and y in R^n by either $x^T y$ or $\langle x, y\rangle$. Vector inequalities are interpreted componentwise. For a set $E \subseteq R^n$, $co\, E$ denotes the convex hull of E and $\overline{co}\, E$ denotes the closure of $co\, E$. For a differentiable function $f : R^n \to R^m$, $\nabla f(\bar{x})$ denotes the Jacobian matrix of f at $\bar{x}$. For a matrix A, A_i denotes the ith row of A.

A function $\phi : R^2 \to R$ is called an NCP function if $\phi(a, b) = 0 \Leftrightarrow ab = 0, a \geq 0, b \geq 0$. For the problem NCP($f$), we define

$$\Phi(x) = \begin{bmatrix} \phi(x_1, f_1(x)) \\ \vdots \\ \phi(x_i, f_i(x)) \\ \vdots \\ \phi(x_n, f_n(x)) \end{bmatrix} \tag{2}$$

and, by abuse of language, call $\Phi(x)$ an NCP function for NCP(f).

We now recall the following definition and examples from Gowda and Ravindran [10].

Definition 1. Given a function $f : \Omega \subseteq R^n \to R^m$ where Ω is an open set in R^n and $x^* \in \Omega$, we say that a nonempty subset $T(x^*)$ (also denoted by $T_f(x^*)$) of $R^{m\times n}$ is an *H-differential of* f at x^* if for every sequence $\{x^k\} \subseteq \Omega$ converging to x^*, there exist a subsequence $\{x^{k_j}\}$ and a matrix $A \in T(x^*)$ such that

$$f(x^{k_j}) - f(x^*) - A(x^{k_j} - x^*) = o\big(\|x_j^k - x^*\|\big). \tag{3}$$

We say that f is H-differentiable at x^* if f has an H-differential at x^*.

A useful equivalent definition of an H-differential $T(x^*)$ is: For any sequence $x^k := x^* + t_k d^k$ with $t_k \downarrow 0$ and $\|d^k\| = 1$ for all k, there exist convergent subsequences $t_{k_j} \downarrow 0$

and $d^{k_j} \to d$, and $A \in T(x^*)$ such that

$$\lim_{j\to\infty} \frac{f\left(x^* + t_{k_j} d^{k_j}\right) - f(x^*)}{t_{k_j}} = Ad.$$

Remarks. As noted by a referee, it is easily seen that if a function $f : \Omega \subseteq R^n \to R^m$ is H-differentiable at a point $\bar{x}$, then there exist a constant $L > 0$ and a neighbourhood $B(\bar{x}, \delta)$ of $\bar{x}$ with

$$\| f(x) - f(\bar{x})\| \le L\|x - \bar{x}\|, \quad \forall x \in B(\bar{x}, \delta). \tag{4}$$

Conversely, if condition (4) holds, then $T(\bar{x}) := R^{m\times n}$ can be taken as an H-differential of f at $\bar{x}$. We thus have, in (4), an alternate description of H-differentiability. But, as we see in the sequel, it is the identification of an appropriate H-differential that becomes important and relevant.

Clearly any function locally Lipschitzian at $\bar{x}$ will satisfy (4). For real valued functions, condition (4) is known as the 'calmness' of f at $\bar{x}$. This concept has been well studied in the literature of nonsmooth analysis (see [24], Chapter 8).

As noted in [10], (i) any superset of an H-differential is an H-differential, (ii) H-differentiability implies continuity, and (iii) H-differentials enjoy simple sum, product and chain rules.

We include the following examples from [10].

Example 1. If $f : R^n \to R^m$ is Fréchet differentiable at $x^* \in R^n$, then f is H-differentiable with $\{\nabla f(x^*)\}$ as an H-differential.

Example 2. Let $f : \Omega \subseteq R^n \to R^m$ be locally Lipschitzian at each point of an open set Ω. Let Ω_f be the set of all points in Ω where f is Fréchet differentiable. For $x^* \in \Omega$, let

$$\partial_B f(x^*) = \{\lim \nabla f(x^k) : x^k \to x^*, x^k \in \Omega_f\}$$

denote the Bouligand subdifferential of f at x^*. Then, the (Clarke) generalized Jacobian [2]

$$\partial f(x^*) = co\, \partial_B f(x^*)$$

is an H-differential of f at x^*.

Example 3. Consider a locally Lipschitzian function $f : \Omega \subseteq R^n \to R^m$ that is semismooth at $x^* \in \Omega$ [17, 20, 22]. This means for any sequence $x^k \to x^*$, and for any $V_k \in \partial f(x^k)$,

$$f(x^k) - f(x^*) - V_k(x^k - x^*) = o(\|x^k - x^*\|).$$

Then the Bouligand subdifferential

$$\partial_B f(x^*) = \{\lim \nabla f(x^k) : x^k \to x^*, x^k \in \Omega_f\}$$

is an H-differential of f at x^*. In particular, this holds if f is piecewise smooth, i.e., there exist continuously differentiable functions $f_j : R^n \to R^m$ such that

$$f(x) \in \{f_1(x), f_2(x), \dots, f_J(x)\} \quad \forall x \in R^n.$$

Example 4. Let $f : R^n \to R^n$ be C-differentiable in a neighborhood D of x^*. This means that there is a compact upper semicontinuous multivalued mapping $x \mapsto T(x)$ with $x \in D$ and $T(x) \subset R^{n\times n}$ satisfying the following condition at any $a \in D$: For $V \in T(x)$,

$$f(x) - f(a) - V(x - a) = o(\|x - a\|).$$

Then, f is H-differentiable at x^* with $T(x^*)$ as an H-differential. See [21] for further details on C-differentiability.

We recall the definitions of $\mathbf{P}_0$ and $\mathbf{P}$-functions (matrices).

Definition 2. For a function $f : R^n \to R^n$, we say that f is a $\mathbf{P}_0(\mathbf{P})$-function if, for any $x \neq y$ in R^n,

$$\max_{\{i:x_i \neq y_i\}} (x - y)_i [f(x) - f(y)]_i \geq 0 \ (> 0). \tag{5}$$

A matrix $M \in R^{n\times n}$ is said to be a $\mathbf{P}_0(\mathbf{P})$-matrix if the function $f(x) = Mx$ is a $\mathbf{P}_0(\mathbf{P})$-function or equivalently, every principle minor of M is nonnegative (respectively, positive [3]).

We note that every monotone (strictly monotone) function is a $\mathbf{P}_0(\mathbf{P})$-function.

The following result is from [18] and [25].

Theorem 1. *Under each the following conditions, $f : R^n \to R^n$ is a $\mathbf{P}_0(\mathbf{P})$-function.*

(a) *f is Fréchet differentiable on R^n and for every $x \in R^n$, the Jacobian matrix $\nabla f(x)$ is a $\mathbf{P}_0(\mathbf{P})$-matrix.*
(b) *f is locally Lipschitzian on R^n and for every $x \in R^n$, the generalized Jacobian $\partial f(x)$ consists of $\mathbf{P}_0(\mathbf{P})$-matrices.*
(c) *f is semismooth on R^n (in particular, piecewise affine or piecewise smooth) and for every $x \in R^n$, the Bouligand subdifferential $\partial_B f(x)$ consists of $\mathbf{P}_0(\mathbf{P})$-matrices.*
(d) *f is H-differentiable on R^n and for every $x \in R^n$, an H-differential $T_f(x)$ consists of $\mathbf{P}_0(\mathbf{P})$-matrices.*

3. Necessary optimality conditions in *H*-differentiable optimization

In this section, we derive necessary optimality conditions for optimization problems involving H-differentiable functions. We first consider the H-differentiability of minimum/maximum of several H-differentiable functions.

Theorem 2. *For $i = 1, \ldots, m$, let $f^i : R^n \to R$ be H-differentiable at x^* with an H-differential $T_{f^i}(x^*)$. Let $f : R^n \to R$ be defined by*

$$f(x) := \min\{f^1(x),\ f^2(x), \ldots,\ f^m(x)\}. \tag{6}$$

Define

$$T_f(x^*) = \cup_{i \in I(x^*)} T_{f^i}(x^*), \tag{7}$$

where $I(x^*) = \{i : f(x^*) = f^i(x^*)\}$. Then f is H-differentiable at x^* with $T_f(x^*)$ as an H-differential. Also, a similar statement holds if 'min' in (6) *is replaced by 'max'.*

Proof: We prove the result for the min-function; the proof of the max-function is similar. Consider a sequence $\{x^k\}$ converging to x^* in R^n. Then there exist $l \in \{1, \ldots, m\}$ and a subsequence $\{x^{k_j}\}$ such that $f(x^{k_j}) = f^l(x^{k_j})$ for all $j = 1, \ldots, \infty$. We have $f(x^*) = f^l(x^*)$ (by the continuity of f^l and f). Now because of the H-differentiability of f^l at x^*, there is a subsequence of $\{x^{k_j}\}$, which we continue to write as $\{x^{k_j}\}$ for simplicity, and a matrix $A^l \in T_{f^l}(x^*)$ such that

$$f^l(x^{k_j}) - f^l(x^*) - A^l(x^{k_j} - x^*) = o(\|x^{k_j} - x^*\|)$$

which leads to

$$f(x^{k_j}) - f(x^*) - A^l(x^{k_j} - x^*) = o(\|x^{k_j} - x^*\|).$$

Since $A^l \in T_{f^l}(x^*) \subseteq \cup_{i \in I(x^*)} T_{f^i}(x^*)$, we see that f is H-differentiable at x^* with $T_f(x^*)$ (defined in (7)) as an H-differential. This completes the proof. □

Remark. In the above theorem, we considered real valued functions. With obvious modifications, one can consider vector valued functions. See Example 8 for an illustration.

Theorem 3. *Suppose $f : R^n \to R$ and x^* is a local optimal solution of the problem*

$$\min_{x \in R^n} f(x).$$

If f is H-differentiable at x^* and $T(x^*)$ is any H-differential, then

$$0 \in \overline{co}\, T(x^*).$$

Proof: Suppose, if possible, that $0 \notin \overline{co}\, T(x^*)$. Since $\overline{co}\, T(x^*)$ is closed and convex, by the strict separation theorem (see p. 50, [15]), there exists a nonzero vector d in R^n such that $Ad < 0$ for all $A \in \overline{co}\, T(x^*)$. From the H-differentiability of f, for the sequence $\{x^* + \frac{1}{k}d\}$, there exist a subsequence $\{x^* + \frac{1}{k_j}d\}$ and $\bar{A} \in T(x^*)$ such that

$$k_j \left[f\left(x^* + \frac{1}{k_j}d\right) - f(x^*) \right] \to \bar{A}d.$$

Since $f(x) \geq f(x^*)$ for all x near x^*, we see that $\bar{A}d \geq 0$ reaching a contradiction. Hence $0 \in \overline{co}\, T(x^*)$. $\square$

Remarks. When f is differentiable at x^* with $T(x^*) = \{\nabla f(x^*)\}$, the above optimality condition reduces to the familiar condition $\nabla f(x^*) = 0$. When f is locally Lipschitzian at $\bar{x}$, the above result reduces to Proposition 2.3.2 in [2] that $0 \in \partial f(x^*)$; see also, Theorem 7 in [17].

The above theorem motivates us to define a *stationary point* of the problem min $f(x)$ as a point x^* such that $0 \in \overline{co}\, T_f(x^*)$ where $T_f(x^*)$ is an H-differential of f at x^*. By weakening this condition, we may call a point x^* a *quasi-stationary point* (*semi-stationary point*) of the problem min $f(x)$ if $0 \in T_f(x^*)$ (respectively, $0 \in co\, T_f(x^*)$). While local/global minimizers of min $f(x)$ are stationary points, it is not clear how to get or describe semi- and quasi- stationary points. However, as we shall see in Sections 6, 7, and 8, they are used in formulating conditions for a point x^* to be a solution of a nonlinear complementarity problem.

We now describe a necessary optimality condition for inequality constrained optimization problems.

Theorem 4. *Suppose that f and g^i $(i = 1, 2 \ldots, m)$ are real valued functions defined on R^n and x^* is a local optimal solution of the problem*

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \quad g^i(x) \leq 0 \quad \textit{for} \quad i = 1, \ldots, m. \end{aligned} \tag{8}$$

Suppose that f and g^i $(i = 1, 2 \ldots, m)$ are H-differentiable at x^* with H-differentials given, respectively, by $T_f(x^*)$ and $T_{g^i}(x^*)$ $(i = 1, 2, \ldots, m)$. Let $g(x) := \max_{1 \leq i \leq m} g^i(x)$ and $I(x^*) = \{i : g(x^*) = g^i(x^*)\}$. Then

$$0 \in \overline{co}\{T_f(x^*) \cup (\cup_{i \in I(x^*)} T_{g^i}(x^*))\}. \tag{9}$$

Proof: We see that x^* is a local optimal solution of the problem

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \quad g(x) \leq 0. \end{aligned} \tag{10}$$

From Theorem 2, we see that g is H-differentiable with $T_g(x^*) := \cup_{i \in I(x^*)} T_{g^i}(x^*)$ as an H-differential. We have to show that $0 \in \overline{co}[T_f(x^*) \cup T_g(x^*)]$. Suppose this statement is false. Then by the strict separation theorem (see p. 50, [15]), there exists a nonzero vector d in R^n such that $Ad < 0$ for all $A \in T_f(x^*) \cup T_g(x^*)$. From the H-differentiability of f and g, for the sequence $\{x^* + \frac{1}{k}d\}$, there exist a subsequence $\{x^* + \frac{1}{k_j}d\}$, matrices $\bar{A} \in T_f(x^*)$ and $\bar{B} \in T_g(x^*)$ such that

$$k_j \left[f\left(x^* + \frac{1}{k_j}d\right) - f(x^*) \right] \to \bar{A}d$$

and

$$k_j \left[g\left(x^* + \frac{1}{k_j} d \right) - g(x^*) \right] \to \bar{B}d.$$

From $\bar{A}d < 0$ and $\bar{B}d < 0$, we see that $f(x^* + \frac{1}{k_j}d) - f(x^*) < 0$ and $g(x^* + \frac{1}{k_j}d) - g(x^*) < 0$ for all large k_j. We reach a contradiction since x^* is assumed to be locally optimal to the given problem. Thus we have the stated conclusion. □

4. *H*-differentials of some NCP functions associated with *H*-differentiable functions

In this section, we describe the H-differentials of some well known NCP functions.

Example 5. Suppose $f : R^n \to R^n$ has an H-differential $T(\bar{x})$ at $\bar{x} \in R^n$. Consider the associated Fischer-Burmeister function [7]

$$\Phi_F(x) := x + f(x) - \sqrt{x^2 + f(x)^2}$$

where all the operations are performed componentwise. Let

$$J(\bar{x}) = \{i : f_i(\bar{x}) = 0 = \bar{x}_i\}.$$

Consider the set Γ of all quadruples (A, V, W, d) with $A \in T(\bar{x})$, $\|d\| = 1$, $V = diag(v_i)$ and $W = diag(w_i)$ are diagonal matrices satisfying the conditions

$$(1 - v_i)^2 + (1 - w_i)^2 = 1 \quad \forall i = 1, 2 \ldots, n \tag{11}$$

and

$$v_i = \begin{cases} 1 - \dfrac{f_i(\bar{x})}{\sqrt{\bar{x}_i^2 + (f_i(\bar{x}))^2}} & \text{when } i \notin J(\bar{x}) \\ 1 - \dfrac{A_i d}{\sqrt{d_i^2 + (A_i d)^2}} & \text{when } i \in J(\bar{x}) \text{ and } d_i^2 + (A_i d)^2 > 0 \\ \text{arbitrary} & \text{when } i \in J(\bar{x}) \text{ and } d_i^2 + (A_i d)^2 = 0, \end{cases}$$

$$w_i = \begin{cases} 1 - \dfrac{\bar{x}_i}{\sqrt{\bar{x}_i^2 + (f_i(\bar{x}))^2}} & \text{when } i \notin J(\bar{x}) \\ 1 - \dfrac{d_i}{\sqrt{d_i^2 + (A_i d)^2}} & \text{when } i \in J(\bar{x}) \text{ and } d_i^2 + (A_i d)^2 > 0 \\ \text{arbitrary} & \text{when } i \in J(\bar{x}) \text{ and } d_i^2 + (A_i d)^2 = 0. \end{cases} \tag{12}$$

We now *claim that* Φ_F (*or* Φ *for simplicity*) *has an* H*-differential at* $\bar{x}$ *given by*

$$S(\bar{x}) = \{VA + W : (A, V, W, d) \in \Gamma\}.$$

To see this claim, let $\bar{x} + t_k d^k \to \bar{x}$ with $t_k \downarrow 0$ and $\|d^k\| = 1$. By the H-differentiability of f, there exist a subsequence $\{t_{k_j}\}$ of $\{t_k\}$, $d^{k_j} \to d$, and $A \in T(\bar{x})$ such that $f(\bar{x} + t_{k_j} d^{k_j}) - f(\bar{x}) - A(t_{k_j} d^{k_j}) = o(t_{k_j})$. Let, for ease of notation, $y^j := \bar{x} + t_{k_j} d^{k_j}$. With A and d, define V and W satisfying (11) and (12); let $B := VA + W$. We claim that $\Phi(y^j) - \Phi(\bar{x}) - B\,(t_{k_j} d^{k_j}) = o(t_{k_j})$. To see this, we fix an index i and show that $\Phi_i(y^j) - \Phi_i(\bar{x}) - [B\,(t_{k_j} d^{k_j})]_i = o(t_{k_j})$. Without loss of generality, let $i = 1$. We consider two cases:

Case (1): $1 \notin J(\bar{x})$.

In this case we have

$$B_1 = \left\{1 - \frac{f_1(\bar{x})}{\sqrt{\bar{x}_1^2 + (f_1(\bar{x}))^2}}\right\} A_1 + \left\{1 - \frac{\bar{x}_1}{\sqrt{\bar{x}_1^2 + (f_1(\bar{x}))^2}}\right\} {e_1}^T,$$

where ${e_1}^T$ is the first row of the identity matrix and

$$\begin{aligned}
&\Phi_1(y^j) - \Phi_1(\bar{x}) - B_1\, t_{k_j} d^{k_j} \\
&= t_{k_j} d_1^{k_j} + \left(A_1\, t_{k_j} d^{k_j}\right) + o\left(t_{k_j}\right) \\
&\quad - \left[\sqrt{\left(\bar{x}_1 + t_{k_j} d_1^{k_j}\right)^2 + \left[f_1(\bar{x}) + A_1\, t_{k_j} d^{k_j} + o(t_{k_j})\right]^2} - \sqrt{\bar{x}_1^2 + (f_1(\bar{x}))^2}\right] \\
&\quad - \left\{1 - \frac{f_1(\bar{x})}{\sqrt{\bar{x}_1^2 + (f_1(\bar{x}))^2}}\right\} A_1\, t_{k_j} d^{k_j} - \left\{1 - \frac{\bar{x}_1}{\sqrt{\bar{x}_1^2 + (f_1(\bar{x}))^2}}\right\} t_{k_j} d_1^{k_j} \\
&= o\left(t_{k_j}\right).
\end{aligned} \tag{13}$$

Case (2): $1 \in J(\bar{x})$.

Subcase (1): $d_1^2 + (A_1 d)^2 > 0$.

In this case,

$$B_1 = \left\{1 - \frac{A_1 d}{\sqrt{d_1^2 + (A_1 d)^2}}\right\} A_1 + \left\{1 - \frac{d_1}{\sqrt{d_1^2 + (A_1 d)^2}}\right\} {e_1}^T,$$

and an easy calculation shows

$$\begin{aligned}
&\Phi_1(y^j) - \Phi_1(\bar{x}) - B_1\, t_{k_j} d^{k_j} \\
&= t_{k_j} d_1^{k_j} + \left(A_1\, t_{k_j} d^{k_j}\right) - \sqrt{\left(t_{k_j} d_1^{k_j}\right)^2 + \left(A_1\, t_{k_j} d^{k_j} + o\left(t_{k_j}\right)\right)^2}
\end{aligned}$$

$$
\begin{aligned}
&- \left\{1 - \frac{A_1 d}{\sqrt{d_1^2 + (A_1 d)^2}}\right\} A_1\, t_{k_j} d^{k_j} \\
&- \left\{1 - \frac{d_1}{\sqrt{d_1^2 + (A_1 d)^2}}\right\} t_{k_j} d_1^{k_j} + o\big(t_{k_j}\big) \\
= {}& o\big(t_{k_j}\big). \qquad (14)
\end{aligned}
$$

Subcase (2): Let $d_1^2 + (A_1 d)^2 = 0$.
In this case $d_1 = 0 = A_1 d$. Then $\Phi_1(y^j) - \Phi_1(\bar{x}) = o(t_{k_j})$.

These arguments prove that $\Phi_i(y^j) - \Phi_i(\bar{x}) - [B\ (t_{k_j} d^{k_j})]_i = o(t_{k_j})$ holds for all i. Thus we have the H-differentiability of Φ with $S(\bar{x})$ as an H-differential.

Remarks. We observe that in the above example, if $T(\bar{x})$ consists of **P**- matrices then $S(\bar{x})$ consists of **P**-matrices. To see this, suppose that every $A \in T(\bar{x})$ is a **P**-matrix and consider any $B = VA + W \in S(\bar{x})$. Since A is a **P**-matrix, there exists an index j with $x_j \neq 0$ such that $x_j[Ax]_j > 0$. Since v_i and w_i in (12) are nonnegative and their sum is positive, $x_j[Bx]_j = x_j[(VA + W)x]_j = v_j[x_j(Ax)_j] + w_j x_j^2 > 0$. It follows that B is a **P**-matrix.

This observation together with Theorem 1 says that if $T(\bar{x})$ consists of **P**-matrices $\forall \bar{x} \in R^n$, then the function Φ_F is a **P**-function. (In fact, Φ_F is a **P**-function whenever f is a continuous **P**-function, see [23].)

We note that $S(\bar{x})$ may not consist of **P**-matrices if f is merely a **P**-function on R^n. This can be seen by the following example. Let $f(x) = x^3$ on R. Then f is a **P**-function and $\Phi_F(x) = x + x^3 - \sqrt{x^6 + x^2}$. By a simple calculation, we see that the $\{2, 0\}$ is an H-differential of Φ_F at zero and that it contains a singular object.

Example 6. In the previous example, we described the H-differential of Fischer-Burmeister function. A similar analysis can be carried out for the NCP function [13]

$$\Phi(x) := x + f(x) - \sqrt{(x - f(x))^2 + \lambda x f(x)} \qquad (15)$$

where λ is a fixed parameter in $(0, 4)$. We note that when $\lambda = 2$, $\Phi(x)$ reduces to the Fischer-Burmeister function, while as $\lambda \to 0$, $\Phi(x)$ becomes

$$\Phi(x) := x + f(x) - \sqrt{(x - f(x))^2}\ (= 2\min\{x,\ f(x)\}).$$

Let

$$J(\bar{x}) = \{i : f_i(\bar{x}) = 0 = \bar{x}_i\}.$$

An H-differential of Φ in (15) is given by

$$S_\lambda(\bar{x}) = \{VA + W : (A, V, W, d) \in \Gamma\},$$

where Γ is the set of all quadruples (A, V, W, d) with $A \in T(\bar{x})$, $\|d\| = 1$, $V = diag(v_i)$ and $W = diag(w_i)$ are diagonal matrices satisfying the conditions

$$(1 - v_i)^2 + (1 - w_i)^2 \in (0, 2) \quad \forall\, i = 1, 2 \ldots, n \tag{16}$$

and

$$v_i = \begin{cases} 1 - \dfrac{-2(\bar{x}_i - f_i(\bar{x})) + \lambda \bar{x}_i}{2\sqrt{(\bar{x}_i - f_i(\bar{x}))^2 + \lambda \bar{x}_i f_i(\bar{x})}} & \text{when } i \notin J(\bar{x}) \\ 1 - \dfrac{-2(d_i - A_i d) + \lambda d_i}{2\sqrt{(d_i - A_i d)^2 + \lambda d_i (A_i d)}} & \text{when } i \in J(\bar{x}) \text{ and } (d_i - A_i d)^2 \\ & \quad + \lambda d_i (A_i d) > 0 \\ \text{arbitrary} & \text{when } i \in J(\bar{x}) \text{ and } (d_i - A_i d)^2 \\ & \quad + \lambda d_i (A_i d) = 0, \end{cases}$$
$$\tag{17}$$
$$w_i = \begin{cases} 1 - \dfrac{2(\bar{x}_i - f_i(\bar{x})) + \lambda f_i(\bar{x})}{2\sqrt{(\bar{x}_i - f_i(\bar{x}))^2 + \lambda \bar{x}_i f_i(\bar{x})}} & \text{when } i \notin J(\bar{x}) \\ 1 - \dfrac{2(d_i - A_i d) + \lambda A_i d}{2\sqrt{(d_i - A_i d)^2 + \lambda d_i (A_i d)}} & \text{when } i \in J(\bar{x}) \text{ and } (d_i - A_i d)^2 \\ & \quad + \lambda d_i (A_i d) > 0 \\ \text{arbitrary} & \text{when } i \in J(\bar{x}) \text{ and } (d_i - A_i d)^2 \\ & \quad + \lambda d_i (A_i d) = 0. \end{cases}$$

Example 7. The following NCP function is called the penalized Fischer-Burmeister function [1]

$$\Phi_\lambda(x) := \lambda \Phi_F(x) + (1 - \lambda) x_+ \, f(x)_+ \tag{18}$$

where $x_+ = \max\{0, x\}$ and $\lambda \in (0, 1)$ is a fixed parameter. Let

$$J(\bar{x}) = \{i : f_i(\bar{x}) = 0 = \bar{x}_i\} \quad \text{and} \quad K(\bar{x}) = \{i : \bar{x}_i > 0,\ f_i(\bar{x}) > 0\}.$$

For Φ_λ in (18), a straightforward calculation shows that an H-differential is given by

$$S(\bar{x}) = \{VA + W : (A, V, W, d) \in \Gamma\},$$

where Γ is the set of all quadruples (A, V, W, d) with $A \in T(\bar{x})$, $\|d\| = 1$, $V = diag(v_i)$ and $W = diag(w_i)$ are diagonal matrices with

$$v_i = \begin{cases} \lambda\left(1 - \dfrac{f_i(\bar{x})}{\sqrt{\bar{x}_i^2 + f_i(\bar{x})^2}}\right) + (1-\lambda)\bar{x}_i & \text{when } i \in K(\bar{x}) \\ \lambda\left(1 - \dfrac{A_i d}{\sqrt{d_i^2 + (A_i d)^2}}\right) & \text{when } i \in J(\bar{x}) \text{ and } d_i^2 + (A_i d)^2 > 0 \\ \lambda\left(1 - \dfrac{f_i(\bar{x})}{\sqrt{\bar{x}_i^2 + f_i(\bar{x})^2}}\right) & \text{when } i \notin J(\bar{x}) \cup K(\bar{x}) \\ \text{arbitrary} & \text{when } i \in J(\bar{x}) \text{ and } d_i^2 + (A_i d)^2 = 0, \end{cases} \tag{19}$$

$$w_i = \begin{cases} \lambda\left(1 - \dfrac{\bar{x}_i}{\sqrt{\bar{x}_i^2 + f_i(\bar{x})^2}}\right) + (1-\lambda) f_i(\bar{x}) & \text{when } i \in K(\bar{x}) \\ \lambda\left(1 - \dfrac{d_i}{\sqrt{d_i^2 + (A_i d)^2}}\right) & \text{when } i \in J(\bar{x}) \text{ and } d_i^2 + (A_i d)^2 > 0 \\ \lambda\left(1 - \dfrac{\bar{x}_i}{\sqrt{\bar{x}_i^2 + f_i(\bar{x})^2}}\right) & \text{when } i \notin J(\bar{x}) \cup K(\bar{x}) \\ \text{arbitrary} & \text{when } i \in J(\bar{x}) \text{ and } d_i^2 + (A_i d)^2 = 0. \end{cases}$$

The above calculation relies on the observation that the following is an H-differential of the one variable function $t \mapsto t_+$ at any $\bar{t}$:

$$\Delta(\bar{t}) = \begin{cases} \{1\} & \text{if } \bar{t} > 0 \\ \{0, 1\} & \text{if } \bar{t} = 0 \\ \{0\} & \text{if } \bar{t} < 0. \end{cases}$$

Example 8. For an H-differentiable function $f : R^n \to R^n$, consider the NCP function

$$\Phi(x) = \min\{x, f(x)\}. \tag{20}$$

We claim that the H-differential of Φ is given by

$$T_\Phi(\bar{x}) = \{VA + W : V = diag(v_i), W = diag(w_i), \text{ with } v_i, w_i \in \{0, 1\}, \\ V + W = I, A \in T_f(\bar{x})\}. \tag{21}$$

To see this claim, let $x^k \to \bar{x}$. By the H-differentiability of f, there exist a subsequence of $\{x^k\}$, which we continue to write as $\{x^k\}$ for simplicity, and a matrix $A \in T_f(\bar{x})$ such that $f(x^k) - f(\bar{x}) - A\,(x^k - \bar{x}) = o(\|x^k - \bar{x}\|)$. By considering a suitable subsequence, if necessary, we may write $\{1, \ldots, n\}$ as a disjoint union of sets α and β where

$$\alpha = \{i : \Phi_i(x^k) = f_i(x^k)\ \forall k\} \quad \text{and} \quad \beta = \left\{i : \Phi_i(x^k) = x_i^k\ \forall k\right\}.$$

Put

$$v_i = \begin{cases} 1 & \text{if } i \in \alpha \\ 0 & \text{if } i \in \beta \end{cases}, \qquad w_i = \begin{cases} 0 & \text{if } i \in \alpha \\ 1 & \text{if } i \in \beta \end{cases},$$
$$V = diag(v_i), \quad W = diag(w_i), \quad \text{and} \quad B := VA + W.$$

We show that $\Phi(x^k) - \Phi(\bar{x}) - B(x^k - \bar{x}) = o(\|x^k - \bar{x}\|)$. To see this, we fix an index j and show that $\Phi_j(x^k) - \Phi_j(\bar{x}) - [B(x^k - \bar{x})]_j = o(\|x^k - \bar{x}\|)$. Let $j = 1$ (for simplicity). We have two cases:

Case (1): $1 \in \alpha$.

$$\begin{aligned}
&[\Phi(x^k) - \Phi(\bar{x}) - (VA + W)(x^k - \bar{x})]_1 \\
&\quad = f_1(x^k) - f_1(\bar{x}) - [VA(x^k - \bar{x})]_1 + [W(x^k - \bar{x})]_1 \\
&\quad = [f(x^k) - f(\bar{x}) - A\,(x^k - \bar{x})]_1 \\
&\quad = o(\|x^k - \bar{x}\|).
\end{aligned}$$

Case (2): $1 \in \beta$. It is easy to verify that $\Phi_1(x^k) - \Phi_1(\bar{x}) - [B(x^k - \bar{x})]_1 = 0 = o(\|x^k - \bar{x}\|)$.

This proves the above claim.

5. The *H*-differentiability of the merit function

In this section, we consider an NCP function Φ corresponding to NCP(f) and let $\Psi := \frac{1}{2}\|\Phi\|^2$.

Theorem 5. *Suppose Φ is H-differentiable at $\bar{x}$ with $S(\bar{x})$ as an H-differential. Then $\Psi := \frac{1}{2}\|\Phi\|^2$ is H-differentiable at $\bar{x}$ with an H-differential given by*

$$T_\Psi(\bar{x}) = \{\Phi(\bar{x})^T B : B \in S(\bar{x})\}.$$

Proof: Consider a sequence $\{\bar{x} + t_k d^k\}$ with $t_k \downarrow 0$ and $\|d^k\| = 1$ for all k. Then there exist $d^{k_j} \to d \in R^n$ and $B \in S(\bar{x})$ such that $\Phi(\bar{x} + t_{k_j} d^{k_j}) - \Phi(\bar{x}) - B(t_{k_j} d^{k_j}) = o(t_{k_j})$. We have

$$\begin{aligned}
\Psi\big(\bar{x} + t_{k_j} d^{k_j}\big) - \Psi(\bar{x}) &= \frac{1}{2}\big\langle \Phi\big(\bar{x} + t_{k_j} d^{k_j}\big), \Phi\big(\bar{x} + t_{k_j} d^{k_j}\big)\big\rangle - \frac{1}{2}\langle \Phi(\bar{x}), \Phi(\bar{x})\rangle \\
&= \frac{1}{2}\big\langle \Phi(\bar{x}) + B\big(t_{k_j} d^{k_j}\big) + o\big(t_{k_j}\big), \Phi(\bar{x}) + B\big(t_{k_j} d^{k_j}\big) + o\big(t_{k_j}\big)\big\rangle \\
&\quad - \frac{1}{2}\langle \Phi(\bar{x}), \Phi(\bar{x})\rangle.
\end{aligned}$$

This gives us

$$\lim_{t_{k_j}\downarrow 0} \frac{\Psi\left(\bar{x}+t_{k_j}d^{k_j}\right)-\Psi(\bar{x})}{t_{k_j}} = \langle \Phi(\bar{x}), Bd\rangle = \Phi(\bar{x})^T Bd.$$

This completes the proof. □

6. Minimizing the merit function under P_0-conditions

For a given function $f : R^n \to R^n$, consider the associated NCP function Φ and the corresponding merit function $\Psi = \frac{1}{2}\|\Phi\|^2$. It should be recalled that

$$\Psi(\bar{x}) = 0 \Leftrightarrow \Phi(\bar{x}) = 0 \Leftrightarrow \bar{x} \text{ solves NCP}(f).$$

One very popular method of finding zeros of Φ is to find the local/global minimum points or 'stationary' points of Ψ. Various researchers have shown, under certain $\mathbf{P}_0$-conditions, that when f is continuously differentiable or more generally locally Lipschitzian, 'stationary' points of Ψ are the zeros of Ψ. In what follows, starting with an H-differentiable function f, we show that under appropriate conditions, a vector $\bar{x}$ is a solution of the NCP(f) if and only if zero belongs to one of the sets $T_\Psi(\bar{x})$, $co\, T_\Psi(\bar{x})$, or $\overline{co}\, T_\Psi(\bar{x})$.

Theorem 6. *Suppose $f : R^n \to R^n$ is H-differentiable at $\bar{x}$ with an H-differential $T(\bar{x})$. Suppose Φ is an NCP function of f. Assume that $\Psi := \frac{1}{2}\|\Phi\|^2$ is H-differentiable at $\bar{x}$ with an H-differential given by*

$$T_\Psi(\bar{x}) = \{\Phi(\bar{x})^T [VA + W] : A \in T(\bar{x}),\ V = diag(v_i),\ \text{and } W = diag(w_i),\ \text{with } v_i\, w_i > 0 \text{ whenever } \Phi_i(\bar{x}) \neq 0\}.$$

Further suppose that $T(\bar{x})$ consists of $\mathbf{P}_0$-matrices. Then

$$0 \in T_\Psi(\bar{x}) \Leftrightarrow \Phi(\bar{x}) = 0.$$

Proof: Clearly, $\Phi(\bar{x}) = 0$ implies that $T_\Psi(\bar{x}) = \{0\}$ by the description of $T_\Psi(\bar{x})$. Conversely, suppose that $0 \in T_\Psi(\bar{x})$, so that for some $\Phi(\bar{x})^T[VA + W] \in T_\Psi(\bar{x})$,

$$0 = \Phi(\bar{x})^T VA + \Phi(\bar{x})^T W$$

yielding $A^T y + z = 0$ where $y = V^T\Phi(\bar{x})$ and $z = W^T\Phi(\bar{x})$. Note that for any index i, $\Phi_i(\bar{x}) \neq 0 \Leftrightarrow y_i \neq 0$ (because $y = V\Phi(\bar{x})$ and $v_i w_i > 0$ when $\Phi_i(\bar{x}) \neq 0$) in which case $y_i(A^T y)_i = -v_i w_i[\Phi_i(\bar{x})]^2 < 0$. Hence if $\Phi(\bar{x}) \neq 0$, then $\max_{\{y_j \neq 0\}} y_j(A^T y)_j < 0$, contradicting the $\mathbf{P}_0$-property of A. We conclude that $\Phi(\bar{x}) = 0$. □

In the next two successive theorems, we replace the condition $0 \in T_\Psi(\bar{x})$ by weaker conditions $0 \in co\, T_\Psi(\bar{x})$ and $0 \in \overline{co}\, T_\Psi(\bar{x})$. Of course, these relaxations come at the expense of imposing either stronger or different conditions on the H-differential of f.

First we recall a definition from [26].

Definition 3. Consider a nonempty set $\mathcal{C}$ in $R^{n\times n}$. We say that a matrix A is a *row representative of* $\mathcal{C}$ if for each index $i = 1, 2, \ldots, n$, the ith row of A is the ith row of some matrix $C \in \mathcal{C}$. We say that $\mathcal{C}$ has the *row-*$\mathbf{P}_0$*-property* (*row-*$\mathbf{P}$*-property*) if every row representative of $\mathcal{C}$ is a $\mathbf{P}_0$-matrix ($\mathbf{P}$-matrix). We say that $\mathcal{C}$ has the *column-*$\mathbf{P}_0$*-property* (*column-*$\mathbf{P}$*-property*) if $\mathcal{C}^T = \{A^T : A \in \mathcal{C}\}$ has the *row-*$\mathbf{P}_0$*-property* (*row-*$\mathbf{P}$*-property*).

We have the following result from [26].

Proposition 1. *A set $\mathcal{C}$ has the row-$\mathbf{P}_0$-property (row-$\mathbf{P}$-property) if and only if for each nonzero x in R^n there is an index i such that $x_i \neq 0$ and $x_i(Cx)_i \geq 0$ (>0) for all $C \in \mathcal{C}$.*

A simple consequence of this proposition is the following.

Corollary 1. *The following statements hold*:

(i) *Suppose the set of matrices $\{A^1, A^2, \ldots, A^L\}$ has the row-$\mathbf{P}_0$-property. Then for any collection $\{V^1, V^2, \ldots, V^L\}$ of nonnegative diagonal matrices, the sum*

$$A^* = \sum_{j=1}^{L} V^j A^j$$

is a $\mathbf{P}_0$-matrix. In particular, any convex combination of the A^is is a $\mathbf{P}_0$-matrix.

(ii) *Suppose the set of matrices $\{A^1, A^2, \ldots, A^L\}$ has the row-$\mathbf{P}$-property. Then for any collection $\{Y^1, \ldots, Y^L, Z^*\}$ of nonnegative diagonal matrices with $Y^1 + \cdots + Y^L + Z^* > 0$,*

$$A^* = \sum_{j=1}^{L} Y^j A^j + Z^*$$

is a $\mathbf{P}$-matrix.

Proof: (i) Let $x \neq 0$ in R^n. By the above proposition, there exists an index i such that $x_i \neq 0$ and $x_i(A^j x)_i \geq 0\ \forall j = 1, \ldots, L$. Clearly $x_i(A^*x)_i = \sum_{j=1}^{L}(V^j)_{ii}[x_i(A^j x)_i] \geq 0$. This proves the $\mathbf{P}_0$-property of A^*. By specializing V^js, we get the additional statement.

(ii) Let $x \neq 0$. By Proposition 1, there exists an index i such that $x_i \neq 0$ and $x_i(A^j x)_i > 0\ \forall j = 1, \ldots, L$. Now we have $x_i(A^*x)_i = \sum_{j=1}^{L}(Y^j)_{ii}\, x_i(A^j x)_i + (Z^*)_{ii} x_i^2$. All the terms of the above sum are nonnegative. If $(Z^*)_{ii} > 0$, then $x_i(A^*x)_i > 0$ (since $x_i^2 > 0$). If $(Z^*)_{ii} = 0$, then $\sum_{j=1}^{L}(Y^j)_{ii} > 0$ which means that $(Y^{j_0})_{ii} > 0$ for some j_0. Since $x_i(A^{j_0}x)_i > 0$, we see that $x_i(A^*x)_i > 0$. Then A^* is a $\mathbf{P}$-matrix. □

Remark. We note that the implications in the above corollary can be reversed: if every A^* in (i) ((ii)) is a $\mathbf{P}_0$-matrix (respectively, $\mathbf{P}$-matrix), then $\{A^1, A^2, \ldots, A^L\}$ has the row-$\mathbf{P}_0$-property (respectively, row-$\mathbf{P}$-property). Peng [19] proves results similar to Corollary 1 under additional/different hypotheses.

Theorem 7. *Suppose $f : R^n \to R^n$ is H-differentiable at $\bar{x}$ with an H-differential $T(\bar{x})$. Suppose that Ψ is H-differentiable at $\bar{x}$ with an H-differential given by*

$$T_\Psi(\bar{x}) = \{\Phi(\bar{x})^T\,[VA + W] : A \in T(\bar{x}),\ V = diag(v_i),\ \text{and } W = diag(w_i), \text{with } v_i \geq 0,\ w_i > 0 \text{ whenever } \Phi_i(\bar{x}) \neq 0\}.$$

Further suppose that $T(\bar{x})$ has the row-$\mathbf{P}_0$-property. Then

$$0 \in co\, T_\Psi(\bar{x}) \Leftrightarrow \Phi(\bar{x}) = 0.$$

Proof: Suppose $\Psi(\bar{x}) = 0$. Then $\Phi(\bar{x}) = 0$ and we have $co\, T_\Psi(\bar{x}) = \{0\}$. Conversely, suppose $0 \in co\, T_\Psi(\bar{x})$. Then by Carathéodory's theorem [15], there exist $\Phi(\bar{x})^T[V^j A^j + W^j] \in T_\Psi(\bar{x})$, and scalars λ_j for $j = 1, 2, \ldots, L$ with $L \leq n + 1$ such that

$$0 = \sum_{i=1}^{L} \Phi(\bar{x})^T \lambda_i [V^i A^i + W^i] \tag{22}$$

where $\sum_{i=1}^{L} \lambda_i = 1,\ \lambda_i > 0\ \ \forall i \in \{1, \ldots, L\}$. We rewrite (22) as

$$0 = \Phi(\bar{x})^T [Y^1 A^1 + \cdots + Y^L A^L + Z^1 + \cdots + Z^L] \tag{23}$$

where $\lambda_i V^i = Y^i$ and $\lambda_i W^i = Z^i$ for all i. Now (23) reduces to

$$0 = (M + Z^*)^T u$$

where $u = \Phi(\bar{x}),\ M = Y^1 A^1 + \cdots + Y^L A^L$, and $Z^* = Z^1 + \cdots + Z^L$. Now writing $|D| = diag(|d_i|)$ for any diagonal matrix $D = diag(d_i)$, we note that $|Y^i|u = Y^i u$ and $|Z^i|u = Z^i u$ for all i. Since the equality $0 = (M + Z^*)^T u$ is unchanged if we replace Y^i by $|Y^i|$ and Z^i by $|Z^i|$, we may assume that Y^i and Z^i are nonnegative for all i. Now suppose, if possible, that $u = \Phi(\bar{x}) \neq 0$. By the above corollary, the matrices M and M^T are $\mathbf{P}_0$-matrices. Therefore, there exists an index i_* such that $u_{i_*} \neq 0$ and $u_{i_*}(M^T u)_{i_*} \geq 0$. From $\Phi(\bar{x})_{i_*} = u_{i_*} \neq 0$, we see that $(W^j)_{i_* i_*} > 0$ and so $(Z^*)_{i_* i_*} > 0$. But

$$0 \leq u_{i_*}(M^T u)_{i_*} = u_{i_*}(-Z^* u)_{i_*} = -(Z^*)_{i_* i_*}\big(u_{i_*}\big)^2$$

is clearly a contradiction since $u_{i_*} \neq 0$. This proves that $\Phi(\bar{x}) = u = 0$. □

Remarks. We note that Theorems 6 and 7 are applicable to the Fischer-Burmeister function $\Phi(x) = \Phi_F(x) = x + f(x) - \sqrt{x^2 + f(x)^2}$. This is because, the set $T_\Psi(\bar{x})$ described in Theorems 6 and 7 is a superset of the H-differential $T_\Psi(\bar{x}) = \{\Phi(\bar{x})^T B : B \in S(\bar{x})\}$ where $S(\bar{x})$ is described in Example 5. (Note that $[\Phi_F(\bar{x})]_i \neq 0 \Rightarrow i \notin J(\bar{x})$ and hence from (12), $v_i,\ w_i > 0$.) Similarly, we see that Theorems 6 and 7 are applicable to the following NCP functions:

- $\Phi(x) = x + f(x) - \sqrt{(x - f(x))^2 + \lambda x f(x)}$. (Clarification Example 6)
- $\Phi(x) = \lambda \Phi_F(x) + (1 - \lambda) x_+ f(x)_+$. (Clarification Example 7)

We state the next result for the Fischer-Burmeister function Φ. However, as in Theorems 6 and 7, it is possible to state a very general result for any NCP function Φ. For simplicity, we avoid dealing in such a generality.

Theorem 8. *Suppose $f : R^n \to R^n$ is H-differentiable at $\bar{x}$ with an H-differential $T(\bar{x})$ which is compact and having the row-$\mathbf{P}_0$-property. Let Φ be the Fischer-Burmeister function as in Example* 5 *and $\Psi := \frac{1}{2}\|\Phi\|^2$. Let $S(\bar{x})$ and $T_\Psi(\bar{x})$ be as in Example* 5 *and Theorem* 5. *Then the following are equivalent*:
(a) *$\bar{x}$ is a local minimizer of Ψ.*
(b) $0 \in \overline{co}\, T_\Psi(\bar{x})$.
(c) $\Phi(\bar{x}) = 0$, *i.e., $\bar{x}$ solves NCP(f).*

Proof: The implication $(a) \Rightarrow (b)$ follows from Theorem 3. The implication $(c) \Rightarrow (a)$ is obvious. We now prove that $(b) \Rightarrow (c)$. Suppose $0 \in \overline{co}\, T_\Psi(\bar{x})$ and assume that $u = \Phi(\bar{x}) \neq 0$. Then there exists a sequence $\{C^k\}$ of matrices in $co\, S(\bar{x})$ such that $0 = \lim u^T C^k$. Now each C^k is a convex combination of at most $n^2 + 1$ matrices of the form $VA + W \in S(\bar{x})$ where $A \in T(\bar{x})$, V and W satisfy (11) and (12). Since $T(\bar{x})$ is compact and the entries of V and W vary over bounded sets in R, we may assume that $C^k \to C$ where C is a convex combination of at most n^2+1 matrices of the form $\bar{V}\,\bar{A} + \bar{W}$ where $\bar{A} \in T(\bar{x})$, $\bar{V}$ and $\bar{W}$ are nonnegative diagonal matrices satisfying a condition like (11) with $\bar{v}_i = 1 - \frac{f_i(\bar{x})}{\sqrt{\bar{x}_i^2 + (f_i(\bar{x}))^2}}$ and $\bar{w}_i = 1 - \frac{\bar{x}_i}{\sqrt{\bar{x}_i^2 + (f_i(\bar{x}))^2}}$ when $i \notin \{f_i(\bar{x}) = 0 = \bar{x}_i\}$. From $0 = \lim u^T C^k$, we get an equation similar to (22) but now with $\bar{V}^i$, $\bar{A}^i$, and $\bar{W}^i$ in place of V^i, A^i, and W^i, respectively. By repeating the argument given in the proof of the previous theorem, we arrive at a contradiction. Hence $\Phi(\bar{x}) = 0$ proving $(b) \Rightarrow (c)$. □

We now state two consequences of the above theorems for the Fischer-Burmeister function (for the sake of simplicity).

Corollary 2. *Let $f : R^n \to R^n$ be differentiable and $\Phi(x)$ be the Fischer-Burmeister function and $\Psi(x) = \frac{1}{2}\|\Phi\|^2$. If f is $\mathbf{P}_0$-function, then $\bar{x}$ is a local minimizer to Ψ if and only if $\bar{x}$ solves NCP(f).*

This corollary is seen from the above theorem by taking $T(\bar{x}) = \{\nabla f(\bar{x})\}$. If we assume the continuous differentiability of f in the above corollary, we get a result of Facchinei and Soares [5]: For a continuously differentiable $\mathbf{P}_0$-function f, every stationary point of Ψ solves NCP(f). (This is because, when f is C^1, Ψ becomes continuously differentiable, see Prop. 3.4 in [5].) See [9] for the monotone case.

Corollary 3. *Let $f : R^n \to R^n$ be locally Lipschitzian. Let Φ be the Fischer-Burmeister function and $\Psi(\bar{x}) = \frac{1}{2}\|\Phi\|^2$. Then the equivalence*

$$0 \in \partial\Psi(\bar{x}) \Leftrightarrow \Psi(\bar{x}) = 0$$

holds under each of the following conditions.
(a) *$\partial f(\bar{x})$ consists of $\mathbf{P}_0$-matrices*;
(b) *$\partial_B f(\bar{x})$ has the row-$\mathbf{P}_0$-property.*

Proof: The stated equivalence under (a) has already been established by Fischer [8]. In fact, by applying Theorem 6 with $T_f(x) = \partial f(x)$ and using his result that $\partial\Psi(x) \subseteq T_\Psi(x)$ for all x, we get the equivalence in (a). Now to see the equivalence under (b), assume (b) holds. Then by Corollary 1, every matrix in $\partial f(\bar{x}) = co\,\partial_B f(\bar{x})$ is a $\mathbf{P}_0$-matrix. Now we have condition (a) and hence the stated equivalence. □

Remark. The condition (b) in the above corollary might be especially useful when the function f is piecewise smooth in which case $\partial_B f(\bar{x})$ consists of a finite number of matrices.

7. Minimizing the merit function under *P*-conditions

The following theorem is similar to Theorem 6.

Theorem 9. *Suppose $f : R^n \to R^n$ is H-differentiable at $\bar{x}$ with an H-differential $T(\bar{x})$. Suppose Φ is an NCP function of f. Assume that $\Psi := \frac{1}{2}\|\Phi\|^2$ is H-differentiable at $\bar{x}$ with an H-differential given by*

$$T_\Psi(\bar{x}) = \{\Phi(\bar{x})^T [VA + W] : A \in T(\bar{x}),\ V = diag(v_i),\ \text{and } W = diag(w_i),\ \text{with } v_i\, w_i \geq 0 \text{ and } v_i + w_i \neq 0 \text{ whenever } \Phi_i(\bar{x}) \neq 0\}.$$

Further suppose that $T(\bar{x})$ consists of **P**-*matrices. Then*

$$0 \in T_\Psi(\bar{x}) \Leftrightarrow \Phi(\bar{x}) = 0.$$

Proof: Suppose $\Phi(\bar{x}) = 0$. Then by description of $T_\Psi(\bar{x})$, we have $T_\Psi(\bar{x}) = \{0\}$. Conversely, suppose that $0 \in T_\Psi(\bar{x})$, so that for some $\Phi(\bar{x})^T [VA + W] \in T_\Psi(\bar{x})$,

$$0 = \Phi(\bar{x})^T VA + \Phi(\bar{x})^T W \Rightarrow A^T y + z = 0$$

where $y = V^T\Phi(\bar{x})$ and $z = W^T\Phi(\bar{x})$. We claim that $\Phi(\bar{x}) = 0$. Suppose, if possible, that $\Phi(\bar{x}) \neq 0$. If $y = 0$, then $z = 0$ which leads to

$$0 = y + z = (V^T + W^T)\,\Phi(\bar{x}) \Rightarrow \Phi_i(\bar{x})\,(V + W)_{ii} = 0\ \forall\, i$$

a contradiction since for some i_0, $\Phi_{i_0}(\bar{x}) \neq 0$ and $v_{i_0} + w_{i_0} \neq 0$. Hence $y \neq 0$ and

$$y_i(A^T y)_i = -z_i y_i = -v_i w_i \Phi_i(\bar{x})^2 \leq 0\ \forall\, i$$

contradicting the **P**-property of A. Hence $\Phi(\bar{x}) = 0$. □

Theorem 10. *Suppose $f : R^n \to R^n$ is H-differentiable at $\bar{x}$ with an H-differential $T(\bar{x})$. Suppose that Ψ is H-differentiable at $\bar{x}$ with an H-differential given by*

$$T_\Psi(\bar{x}) = \{\Phi(\bar{x})^T [VA + W] : A \in T(\bar{x}),\ V = diag(v_i),\ \text{and } W = diag(w_i), \text{with } v_i \geq 0,\ w_i \geq 0,\ \text{and } v_i + w_i \neq 0 \text{ whenever } \Phi_i(\bar{x}) \neq 0\}.$$

*Further suppose that $T(\bar{x})$ has the row-**P**-property. Then*

$$0 \in co\, T_\Psi(\bar{x}) \Leftrightarrow \Phi(\bar{x}) = 0.$$

Proof: The proof is similar to that of Theorem 7. To show that $0 \in co\, T_\Psi(\bar{x}) \Rightarrow \Phi(\bar{x}) = 0$, we proceed as in the proof of Theorem 7. We have statements (22) and (23) in our new setting where we may assume (as before) that Y^i and Z^i are nonnegative for all i. Since $\sum_{i=1}^{L}(Y^i + Z^i) > 0$, . by Corollary 1 (ii), taking $Z^* = \sum_{i=1}^{L} Z^i$, we see that the matrix in (23) is nonsingular. It follows that $\Phi(\bar{x}) = 0$. □

Remark. We note that Theorems 9 and 10 are applicable to the min-function Φ of Example 8.

8. Minimizing the merit function under regularity (strict regularity) conditions

We now generalize the concept of a regular (strictly regular) point [14] in order to weaken the hypotheses in the Theorems 6 and 7.

For a given H-differentiable function f and $\bar{x} \in R^n$, we define the following subsets of $I = \{1, 2, \ldots, n\}$.

$$\mathcal{C}(\bar{x}) := \{i \in I : \bar{x}_i \geq 0,\ f_i(\bar{x}) \geq 0,\ \bar{x}_i f_i(\bar{x}) = 0\}, \quad \mathcal{R}(\bar{x}) := I \backslash \mathcal{C}(\bar{x}),$$
$$\mathcal{P}(\bar{x}) := \{i \in \mathcal{R}(\bar{x}) : \bar{x}_i > 0,\ f_i(\bar{x}) > 0\}, \quad \mathcal{N}(\bar{x}) := \mathcal{R}(\bar{x}) \backslash \mathcal{P}(\bar{x}).$$

Definition 4. Consider f, $\bar{x}$, and the index sets as above. Let $T(\bar{x})$ be an H-differential of f at $\bar{x}$. Then the vector $\bar{x} \in R^n$ is called a *regular* (*strictly regular*) point of f with respect to $T(\bar{x})$ if for every nonzero vector $z \in R^n$ such that

$$z_{\mathcal{C}} = 0, \qquad z_{\mathcal{P}} > 0, \qquad z_{\mathcal{N}} < 0, \tag{24}$$

there exists a vector $s \in R^n$ such that

$$s_{\mathcal{P}} \geq 0, \qquad s_{\mathcal{N}} \leq 0, \qquad s_{\mathcal{R}} \neq 0, \text{ and} \tag{25}$$
$$s^T A^T z \geq 0\ (> 0) \quad \text{for all } A \in T(\bar{x}). \tag{26}$$

Theorem 11. *Suppose $f : R^n \to R^n$ is H-differentiable at $\bar{x}$ with an H-differential $T(\bar{x})$. Let Φ be an NCP function satisfying the following conditions*:

$$\begin{aligned} &i \in \mathcal{P} \Rightarrow \Phi_i(\bar{x}) > 0,\\ &i \in \mathcal{N} \Rightarrow \Phi_i(\bar{x}) < 0,\\ &i \in \mathcal{C} \Rightarrow \Phi_i(\bar{x}) = 0. \end{aligned} \tag{27}$$

Suppose Ψ is H-differentiable with an H-differential given by

$$\begin{aligned} T_\Psi(\bar{x}) = \{&\Phi(\bar{x})^T[VA + W] :\ A \in T(\bar{x}),\ V = diag(v_i),\\ &W = diag(w_i),\ with\ v_i > 0,\ w_i > 0\ whenever\ \Phi(\bar{x})_i \neq 0\}. \end{aligned} \tag{28}$$

Then $0 \in T_\Psi(\bar{x})$ and $\bar{x}$ is a regular point if and only if $\bar{x}$ solves NCP(f).

Proof: Suppose that $0 \in T_\Psi(\bar{x})$ and $\bar{x}$ is a regular point. Then for some $\Phi(\bar{x})^T\,[VA + W] \in T_\Psi(\bar{x})$,

$$0 = \Phi(\bar{x})^T V A + \Phi(\bar{x})^T W \Rightarrow A^T z + y = 0 \tag{29}$$

where $z = V^T \Phi(\bar{x})$ and $y = W^T \Phi(\bar{x})$. For any $s \in R^n$, (29) yields

$$s^T A^T z + s^T y = 0. \tag{30}$$

We claim that $\Phi(\bar{x}) = 0$. Assume the contrary that $\bar{x}$ is not a solution of NCP(f). Then $\mathcal{R} \neq \emptyset$ and $z_{\mathcal{C}} = 0,\ z_{\mathcal{P}} > 0,\ z_{\mathcal{N}} < 0$. Since $\bar{x}$ is a regular point, and y and z have the same sign, by taking a vector $s \in R^n$ satisfying (25) and (26), we have

$$s^T A^T z \geq 0 \tag{31}$$

and

$$s^T y = s_{\mathcal{C}}^T y_{\mathcal{C}} + s_{\mathcal{P}}^T y_{\mathcal{P}} + s_{\mathcal{N}}^T y_{\mathcal{N}} > 0. \tag{32}$$

Clearly (31) and (32) contradict (30). Hence $\bar{x}$ is a solution to NCP(f). The 'if' part of the theorem follows easily from the definitions. □

Remark. Theorem 11 is applicable to the NCP functions of Examples 5, 6 and 7.

A slight modification of the above theorem leads to the following result.

Theorem 12. *Suppose $f : R^n \to R^n$ is H-differentiable at $\bar{x}$ with an H-differential $T(\bar{x})$. Let Φ be an NCP function satisfying the following conditions*:

$$\begin{aligned} i \in \mathcal{P} &\Rightarrow \Phi_i(\bar{x}) > 0,\\ i \in \mathcal{N} &\Rightarrow \Phi_i(\bar{x}) < 0,\\ i \in \mathcal{C} &\Rightarrow \Phi_i(\bar{x}) = 0. \end{aligned} \tag{33}$$

Suppose Ψ is H-differentiable with an H-differential given by

$$\begin{aligned} T_\Psi(\bar{x}) = \{\Phi(\bar{x})^T[VA + W] : A \in T(\bar{x}),\ V = diag(v_i),\\ W = diag(w_i),\ \text{with } v_i > 0,\ w_i \geq 0 \text{ whenever } \Phi(\bar{x})_i \neq 0\}. \end{aligned} \tag{34}$$

Then $0 \in T_\Psi(\bar{x})$ and $\bar{x}$ is a strictly regular point if and only if $\bar{x}$ solves NCP(f).

Proof: The proof is similar to that of Theorem 11. □

Concluding remarks

In this paper, we considered two applications of H-differentiability. The first application dealt with the necessary optimality condition in H-differentiable optimization. In the second application, for a nonlinear complementarity problem corresponding to an H-differentiable function, with an associated NCP function Φ and a merit function $\Psi = \frac{1}{2}\|\Phi\|^2$, we described conditions under which every global/local minimum or a stationary point of Ψ is a solution of NCP(f). We would like to note here that similar methodologies can be carried out for other merit functions. For example, we can consider the Implicit Lagrangian function of Mangasarian and Solodov [16]:

$$\Phi(x) := x * f(x) + \frac{1}{2\alpha}[\max^2\{0, x - \alpha f(x)\} - x^2 + \max^2\{0, f(x) - \alpha x\} - f(x)^2],$$

where $\alpha > 1$ is any fixed parameter and $x * y$ is the Hadamard (=componentwise) product of vectors x and y. (In [16], it is shown that $\Phi(\bar{x}) = 0 \Leftrightarrow \bar{x}$ solves NCP(f).)

By defining the merit function

$$\Psi(\bar{x}) := \sum_{i=1}^{n} \Phi_i(\bar{x})$$

and formulating the concept of strictly regular point, we can extend the results of [4] for H-differentiable functions.

Our results recover/extend various well known results stated for continuously differentiable (locally Lipschitzian, semismooth, C-differentiable) functions.

Acknowledgments

We thank the referees for their constructive comments.

References

1. B. Chen, X. Chen, and C. Kanzow, "A penalized Fischer-Burmeister NCP-function: Theoretical investigation and numerical results," Applied Mathematics Report 97/28, School of Mathematics, The university of New South Wales, Sydney 2052, Australia, Sept. 1997.
2. F.H. Clarke, Optimization and Nonsmooth Analysis, Wiley: New York, 1983 (reprinted by SIAM: Philadelphia, PA, 1990).
3. R.W. Cottle, J.-S. Pang, and R.E. Stone, The Linear Complementarity Problem, Academic Press: Boston, 1992.
4. F. Facchinei and C. Kanzow, "On unconstrained and constrained stationary points of the implicit lagrangian," Journal of Optimization Theory and Applications, vol. 92, pp. 99–115, 1997.
5. F. Facchinei and J. Soares, "A new merit function for nonlinear complementarity problems and related algorithm," SIAM Journal on Optimization, vol. 7, pp. 225–247, 1997.
6. F. Facchinei, A. Fischer, and C. Kanzow, "Regularity properties of a semismooth reformulation of variational inequalities," SIAM Journal on Optimization, vol. 8, pp. 850–869, 1998.
7. A. Fischer, "A special Newton-type optimization method," Optimization, vol. 24, pp. 269–284, 1992.

8. A. Fischer, “Solution of monotone complementarity problems with locally Lipschitzian functions,” Math. Prog., vol. 76, pp. 513–532, 1997.
9. C. Geiger and C. Kanzow, “On the resolution of monotone complementarity problems,” Computational Optimization, vol. 5, pp. 155–173, 1996.
10. M.S. Gowda and G. Ravindran, “Algebraic univalence theorems for nonsmooth functions,” Research Report, Department of Mathematics and Statistics, University of Maryland, Baltimore County, Baltimore, Maryland 21250, March 15, 1998, to appear in Journal of Mathematical Analysis and Applications.
11. H. Jiang and L. Qi, “A new nonsmooth equations approach to nonlinear complementarity problems,” SIAM Journal on Control and Optimization, vol. 35, pp. 178–193, 1997.
12. H. Jiang, “Unconstrained minimization approaches to nonlinear complementarity problems,” Journal of Global Optimization, vol. 9, pp. 169–181, 1996.
13. C. Kanzow and H. Kleinmichel, “A new class of semismooth Newton-type methods for nonlinear complementarity problems,” Computational Optimization and Applications, vol. 11, pp. 227–251, 1998.
14. T.D. Luca, F. Facchinei, and C. Kanzow, “A semismooth equation approach to the solution of nonlinear complementarity problems,” Math. Prog., vol. 75, pp. 407–439, 1996.
15. O.L. Mangasarian, Nonlinear Programming, McGraw-Hill: New York, 1969.
16. O.L. Mangasarian and M.V. Solodov, “Nonlinear complementarity as unconstrained and constrained minimization,” Math. Prog., vol. 62, pp. 277–297, 1993.
17. R. Mifflin, “Semismooth and semiconvex functions in constrained optimization,” SIAM Control and Opt., vol. 15, pp. 952–972, 1977.
18. J.J. Moré and W.C. Rheinboldt, “On P- and S- functions and related classes of N-dimensional nonlinear mappings,” Linear Algebra Appl., vol. 6, pp. 45–68, 1973.
19. J.-M. Peng, “A smoothing function and its applications,” in Reformulation-Nonsmooth, Piecewise, Semismooth and Smoothing Methods, M. Fukushima and L. Qi (Eds.), Kluwer Academic Publisher, Dordrecht, 1999.
20. L. Qi, “Convergence analysis of some algorithms for solving nonsmooth equations,” Math. Oper. Res., vol. 18, pp. 227–244, 1993.
21. L. Qi, “C-differentiability, C-differential operators and generalized Newton methods,” Research Report, School of Mathematics, The University of New South Wales, Sydney, New South Wales 2052, Australia, Jan. 1996.
22. L. Qi and J. Sun, “A nonsmooth version of Newton’s method,” Math. Prog., vol. 58, pp. 353–367, 1993.
23. G. Ravindran and M.S. Gowda, “Regularization of $\mathbf{P}_0$-functions in box variational inequality problems,” Research Report, Department of Mathematics and Statistics, University of Maryland, Baltimore County, Baltimore, Maryland 21250, Aug. 15, 1997, to appear in SIAM Journal on Optimization.
24. R.T. Rockafellar and R.J.-B. Wets, Variational Analysis, Grundlehren der Mathematischen Wissenschaften, vol. 317, Springer-Verlag: Berlin, 1998.
25. Y. Song, M.S. Gowda, and G. Ravindran, “On characterizations of **P**- and $\mathbf{P}_0$- properties in nonsmooth functions,” Research Report, Department of Mathematics and Statistics, University of Maryland, Baltimore County, Baltimore, Maryland 21250, June 1998, to appear in Mathematics of Operations Research.
26. Y. Song, M.S. Gowda, and G. Ravindran, “On some properties of **P**-matrix sets,” Linear Algebra and its Applications, vol. 290, pp. 237–246, 1999.