# GENERALIZATIONS OF THE DENNIS-MORÉ THEOREM II[1]

**Asen L. Dontchev**
Mathematical Reviews, Ann Arbor, MI 48107-8604

**Abstract.** This paper is a continuation of our previous paper [3] were we presented generalizations of the Dennis-Moré theorem to characterize q-superliner convergences of quasi-Newton methods for solving equations and variational inequalities in Banach spaces. Here we prove Dennis-Moré type theorems for inexact quasi-Newton methods applied to variational inequalities in finite dimensions. We first consider variational inequalities for functions that are merely Lipschitz continuous. Then we present a parallel result for semismooth functions. An erratum to a theorem in our previous paper is also given.

**Key Words.** inexact quasi-Newton method, variational inequality, semismooth functions, strong metric subregularity, q-superlinear convergence.

**AMS Subject Classification (2010)** 49J53, 49J40, 65J15, 90C33.

The celebrated Dennis-Moré theorem [1] characterizes the $q$-superlinear convergence of quasi-Newton methods of the form

$$f(x_k) + B_k(x_{k+1} - x_k) = 0, \quad k = 0, 1, \ldots, \quad x_0 \text{ given}, \tag{1}$$

for finding a zero of a smooth function $f : \mathbb{R}^n \to \mathbb{R}^n$, where $B_k$ is a sequence of matrices constructed in certain way. Throughout, for a sequence $\{x_k\}$ denote $s_k = x_{k+1} - x_k$ and $e_k = x_k - \bar{x}$. Recall that $x_k \to \bar{x}$ $q$-superlinearly when

$$\lim_{k\to\infty} \frac{\|e_{k+1}\|}{\|e_k\|} = 0.$$

The following statement of the Dennis-Moré theorem is slightly different, though equivalent, to that originally given in [1].

**Theorem 1** (Dennis-Moré [1]). *Suppose that $f : \mathbb{R}^n \to \mathbb{R}^n$ is strictly differentiable at $\bar{x}$, a zero of $f$, and the Jacobian mapping $\nabla f(\bar{x})$ is nonsingular. Let $\{B_k\}$ be a sequence of $n \times n$ matrices, let $E_k = B_k - \nabla f(\bar{x})$, and let the sequence $\{x_k\}$ be generated by (1) and converge to $\bar{x}$. Then $x_k \to \bar{x}$ q-superlinearly if and only if*

$$\lim_{k\to\infty} \frac{\|E_k s_k\|}{\|s_k\|} = 0. \tag{2}$$

[1]This work is supported by the National Science Foundation Grant DMS 1008341 through the University of Michigan.

In a path-breaking paper Dembo, Eisenstat and Steihaug [2] introduced an inexact version of the Newton method of the following form: given a sequence of positive scalars $\eta_k$ and a starting point $x_0$, the $(k+1)$st iterate is chosen to satisfy the condition

$$\|f(x_k) + \nabla f(x_k)(x_{k+1} - x_k)\| \le \eta_k \|f(x_k)\|. \tag{3}$$

In particular, the following result was proved in [2, Theorem 3.4, Corollary 3.5]:

**Theorem 2** (Dembo, Eisenstat and Steihaug [2])**.** *Suppose that $f$ is continuously differentiable in a neighborhood of $\bar{x}$, a zero of $f$, and the Jacobian $\nabla f(\bar{x})$ is nonsingular. Let $\eta_k \to 0$. Consider a sequence $\{x_k\}$ generated by* (3) *which is convergent to $\bar{x}$. Then $x_k \to \bar{x}$ superlinearly.*

Basic information about the inexact method (3) is given in the book of Kelley [7, Chapter 6], where convergence and numerical implementations are discussed. An extension of the work of Dembo et al. [2] for generalized equations is presented in [4].

In a previous paper [3] we presented generalizations of the Dennis-Moré theorem for exact quasi-Newton methods applied to nonsmooth and generalized equations in Banach spaces. In this paper we focus on inexact quasi-Newton methods for variational inequalities in finite dimensions. We present first a theorem which generalizes both theorems 1 and 2 above for functions $f$ which are merely Lipschitz continuous. This result is obtained as a particular case of a more general Dennis-Moré type theorem characterizing q-superliner convergence of inexact quasi-Newton methods applied to variational inequalities involving Lipschitz continuous functions. A related, but different theorem is established for variational inequalities involving semismooth functions. Throughout, $\mathbb{R}^n$ is the $n$-dimensional Euclidean space equipped with the usual norm $\|\cdot\|$, $I\!\!B$ is the unit ball, and $d(x, C)$ denotes the distance from a point $x$ to a set $C$.

In preparation to stating our first result, recall that the Clarke generalized Jacobian $\partial f(x)$ of a function $f : \mathbb{R}^n \to \mathbb{R}^m$ at any $x$ around which $f$ is Lipschitz continuous, is the convex hull of the set consisting of all matrices $A \in \mathbb{R}^{m\times n}$ for which there is a sequence of points $x_k \to x$ such that $f$ is differentiable at $x_k$ and $\nabla f(x_k) \to A$. The set $\partial f(x)$ is a nonempty, convex and compact subset of $\mathbb{R}^{m\times n}$. Furthermore, the mapping $x \mapsto \partial f(x)$ has closed graph and is upper semicontinuous at $x$, meaning that for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\partial f(x') \subset \partial f(x) + \varepsilon I\!\!B_{m\times n} \qquad \text{for all } x' \in I\!\!B_\delta(x).$$

Also recall Clarke's mean value theorem, according to which for any $x', x''$ in an open ball around $x$,

$$f(x') - f(x'') = A(x' - x'') \quad \text{for some} \quad A \in \mathrm{co} \bigcup_{t\in[0,1]} \partial f(tx' + (1-t)x'').$$

A key role in our analysis is played by the following Folk Theorem which can be traced back to [6] if not earlier; we state it as a proposition and supply with a proof for completeness.

**Proposition 3.** *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz continuous around $\bar{x}$. Then for every $\varepsilon > 0$ there exists $\delta > 0$ such that for every $x, x' \in I\!\!B_\delta(\bar{x})$ there exists $A \in \partial f(\bar{x})$ with the property*

$$\|f(x) - f(x') - A(x - x')\| \le \varepsilon \|x - x'\|.$$

**Proof.** Let $\varepsilon > 0$. From the upper semicontinuity of $\partial f$ there exists $\delta > 0$ such that

$$\partial f(x) \subset \partial f(\bar{x}) + \varepsilon \mathbb{B}_{n\times n} \quad \text{for all} \quad x \in \mathbb{B}_\delta(\bar{x}).$$

Thus, for any $x, x' \in \mathbb{B}_\delta(\bar{x})$,

$$\partial f(tx + (1-t)x') \subset \partial f(\bar{x}) + \varepsilon \mathbb{B}_{n\times n}.$$

The set on the right side of this inclusion is convex and does not depend on $t$, hence

$$\text{co} \bigcup_{t\in[0,1]} \partial f(tx + (1-t)x') \subset \partial f(\bar{x}) + \varepsilon \mathbb{B}_{n\times n}.$$

But then, from Clarke's mean value theorem, there exists $A \in \partial f(\bar{x})$ with the desired property. $\square$

We will use the following immediate corollary of Proposition 3.

**Corollary 4.** *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz continuous around $\bar{x}$ and consider a sequence $x_k \to \bar{x}$. Then there exists a sequence of matrices $A_k \in \partial f(\bar{x})$ such that*

$$\lim_{k\to\infty} \frac{\|f(x_{k+1}) - f(x_k) - A_k s_k\|}{\|s_k\|} = 0. \tag{4}$$

We consider first the following inexact quasi-Newton method:

$$\|f(x_k) + B_k(x_{k+1}) - x_k)\| \le \eta_k \|f(x_k)\|, \tag{5}$$

where $\{B_k\}$ is a sequence of $n \times n$ matrices and $\eta_k$ is a sequence of positive numbers. Our first result generalizes both theorems 1 and 2.

**Theorem 5.** *Consider a function $f : \mathbb{R}^n \to \mathbb{R}^n$ with a zero $\bar{x}$ which is Lipschitz continuous in a neighborhood $U$ of $\bar{x}$ and there exists $\kappa > 0$ such that*

$$\|x - \bar{x}\| \le \kappa \|f(x)\| \quad \textit{for all } x \in U. \tag{6}$$

*Let $\{B_k\}$ be a sequence of $n\times n$ matrices and let $\eta_k \to 0$. Consider a sequence $\{x_k\}$ generated by (5) which is convergent to $\bar{x}$, and an associated sequence of matrices $A_k \in \partial f(\bar{x})$ satisfying (4) whose existence is claimed in Corollary 4. Let $E_k = B_k - A_k$. Then $x_k \to \bar{x}$ q-superlinearly if and only if*

$$\lim_{k\to\infty} \frac{\|E_k s_k\|}{\|s_k\|} = 0.$$

We will obtain Theorem 5 as a particular case of a more general result concerning the variational inequality

$$f(x) + N_C(x) \ni 0, \tag{7}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ and $N_C$ is the normal cone mapping to a convex polyhedral set $C \subset \mathbb{R}^n$, defined as

$$N_C(x) = \begin{cases} \{y \mid \langle y, v - x\rangle \le 0 \ \text{ for all } \ v \in C\} & \text{for } x \in C, \\ \emptyset & \text{otherwise.} \end{cases}$$

Let $\Pi_C$ be the Euclidean projector on $C$. Recall that solving the variational inequality (7) is equivalent to solving the equation

$$\varphi(x) = 0 \quad \text{where} \quad \varphi(x) = \Pi_C(x - f(x)) - x. \tag{8}$$

We consider the following inexact quasi-Newton method for solving (7):

$$d(0, f(x_k) + B_k s_k + N_C(x_{k+1})) \leq \eta_k \|\varphi(x_k)\|, \tag{9}$$

where $\varphi$ is defined in (8). When $C = \mathbb{R}^n$ then the iteration (9) reduces to (5). Denote by $K$ the critical cone to $C$ at $\bar{x}$ for $-f(\bar{x})$, that is $K = \{w \in T_C(\bar{x}) \mid w \perp f(\bar{x})\}$, where $T_C(x)$ is the tangent cone to the set $C$ at $x$.

In further lines we employ the concept of strong subregularity. A generally set-valued mapping $H : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is said to be *strongly subregular* at $\bar{x}$ for $\bar{y}$ when $\bar{y} \in H(\bar{x})$ and there is a constant $\kappa > 0$ together with a neighborhood $U$ of $\bar{x}$ such that

$$\|x - \bar{x}\| \leq \kappa d(\bar{y}, H(x)) \quad \text{for all} \quad x \in U. \tag{10}$$

In particular, when the mapping $H$ is a function $f$, condition (10) becomes (6). Strong subregularity of a mapping $H$ at $\bar{x}$ for $\bar{y}$ implies that $\bar{x}$ is an isolated point in $H^{-1}(\bar{y})$; moreover, it is equivalent to the so-called isolated calmness property of the inverse mapping $H^{-1}$. The property of strong subregularity obeys the general paradigm of the inverse function theorem: of a stronly subregular mapping $H$ is perturbed by adding a function $f$ with a sufficiently small Lipschitz constant, then the sum $H + f$ remains strongly subregular. In particular, if a function $f$ is strictly differentiable at $\bar{x}$ then for any set-valued mapping $F$ the mapping $f + F$ is strongly subregular at $\bar{x}$ for $\bar{y}$ if and only if the linearized mapping $x \mapsto f(\bar{x}) + \nabla f(\bar{x})(x - \bar{x}) + F(x)$ is strongly subregular at $\bar{x}$ for $\bar{y}$. Another basic fact is that any mapping $H : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, whose graph is the union of finitely many convex polyhedral sets, is strongly subregular at $\bar{x}$ for $\bar{y}$ if and only if $\bar{x}$ is an isolated point in $H^{-1}(\bar{y})$. Combining these two properties we obtain that the mapping $f + N_C$, where $f$ is strictly differentiable at $\bar{x}$ and $C$ is a convex polyhedral, is strongly subregular at $\bar{x}$ for 0 if and only if $\bar{x}$ is an isolated solution of the linearization $x \mapsto f(\bar{x}) + \nabla f(\bar{x})(x - \bar{x}) + N_C(x)$. The latter in turn is equivalent to the condition $(\nabla f(\bar{x}) + N_K)^{-1}(0) = \{0\}$ where $K$ is the critical cone to $C$ at $\bar{x}$ for $-f(\bar{x})$. All this can be found in the book [5] together with a broad discussion of other regularity properties in variational analysis.

The theorem proved next extends Theorem 5 for the method (9).

**Theorem 6.** *Let $\bar{x}$ be a solution of the variational inequality* (7) *and let $f$ be Lipschitz continuous in a neighborhood of $\bar{x}$. Let $\{B_k\}$ be a sequence of $n \times n$ matrices and let $\eta_k \to 0$. Consider a sequence $\{x_k\}$ generated by* (9) *which is convergent to $\bar{x}$. Let $\{A_k\}$ be an associated sequence of matrices $A_k \in \partial f(\bar{x})$ satisfying* (4) *whose existence is claimed in Corollary 4, and let $E_k = B_k - A_k$. If $x_k \to \bar{x}$ q-superlinearly then*

$$\lim_{k \to \infty} \frac{d(0, E_k s_k + N_K(e_{k+1}))}{\|s_k\|} = 0. \tag{11}$$

*Conversely, if the mapping $f + N_C$ is strongly subregular at $\bar{x}$ for 0 and*

$$\lim_{k \to \infty} \frac{\|E_k s_k\|}{\|s_k\|} = 0. \tag{12}$$

*then $x_k \to \bar{x}$ q-superlinearly.*

**Proof.** Let $x_k \to \bar{x}$ $q$-superlinearly and let $\varepsilon > 0$. In [1, Lemma 2.1] it is shown that

$$\frac{\|s_k\|}{\|e_k\|} \to 1 \quad \text{as} \quad k \to \infty. \tag{13}$$

Then, for large $k$ we get

$$\|e_{k+1}\| \le \varepsilon \|s_k\| \qquad \text{and} \qquad \|e_k\| \le 2\|s_k\|. \tag{14}$$

From iteration (9) we obtain that for each $k$ there exists $y_k$ such that

$$y_k \in f(x_k) + B_k s_k + N_C(x_{k+1}) \tag{15}$$

and

$$\|y_k\| \le \eta_k \|\varphi(x_k)\|. \tag{16}$$

Adding and subtracting to the left side of (15) we have

$$f(\bar{x}) - f(\bar{x}) - f(x_{k+1}) + f(x_{k+1}) - f(x_k) - A_k s_k + y_k \in E_k s_k + N_C(\bar{x} + e_{k+1}).$$

Reduction Lemma 2E.4 in [5] implies that

$$f(\bar{x}) - f(x_{k+1}) + f(x_{k+1}) - f(x_k) - A_k s_k + y_k \in E_k s_k + N_K(e_{k+1}). \tag{17}$$

Note that $\varphi(\bar{x}) = 0$ and $\varphi$ is Lipschitz continuous around $\bar{x}$, hence there exists $\ell > 0$ such that, from (16),

$$\|y_k\| \le \eta_k \|\varphi(x_k)\| = \eta_k \|\varphi(x_k) - \varphi(\bar{x})\| \le \eta_k \ell \|e_k\|. \tag{18}$$

Using (14), for sufficiently large $k$,

$$\|y_k\| \le 2\eta_k \ell \|s_k\| \tag{19}$$

Let $l$ be the Lipschitz constant of $f$ near $\bar{x}$; then

$$\|f(\bar{x}) - f(x_{k+1})\| \le l \|e_{k+1}\| \le l\varepsilon \|s_k\|. \tag{20}$$

From (4), for large $k$,

$$\|f(x_{k+1}) - f(x_k) - A_k s_k\| \le \varepsilon \|s_k\|. \tag{21}$$

Using (19), (20) and (21), we obtain

$$\begin{aligned}
&\|f(\bar{x}) - f(x_{k+1}) + f(x_{k+1}) - f(x_k) - A_k s_k + y_k\| \\
&\quad \le \|y_k\| + \|f(\bar{x}) - f(x_{k+1})\| + \|f(x_{k+1}) - f(x_k) - A_k s_k\| \\
&\quad \le 2\eta_k \ell \|s_k\| + l\varepsilon \|s_k\| + \varepsilon \|s_k\|.
\end{aligned}$$

Taking into account (17), this yields

$$d(0, E_k s_k + N_K(e_{k+1})) \le 2\eta_k \ell \|s_k\| + (l+1)\varepsilon \|s_k\|.$$

Since $\eta_k \to 0$ is $\varepsilon$ can be arbitrarily small, we obtain (11).

Now, suppose that the mapping $f + N_C$ is strongly subregular at the solution $\bar{x}$ for 0 and consider a sequence $x_k \to \bar{x}$ generated by (9) for a sequence of matrices $\{B_k\}$. Let $\{A_k\}$ be a sequence of matrices $A_k \in \partial f(\bar{x})$ satisfying (4) and suppose that (12) is satisfied. From the strong subregularity, there exists a constant $\kappa > 0$ such that, for large $k$,

$$\|e_{k+1}\| \leq \kappa d(0, f(x_{k+1}) + N_C(x_{k+1})). \tag{22}$$

As in the beginning of the proof, there exists $y_k$ satisfying (15) and (16); then

$$y_k - f(x_k) - A_k s_k - E_k s_k + f(x_{k+1}) \in f(x_{k+1}) + N_C(x_{k+1}).$$

Hence, from (22),

$$\begin{aligned} \|e_{k+1}\| &\leq \kappa\|y_k - f(x_k) - A_k s_k - E_k s_k + f(x_{k+1})\| \\ &\leq \kappa\|y_k\| + \kappa\|f(x_{k+1}) - f(x_k) - A_k s_k\| + \kappa\|E_k s_k\|. \end{aligned} \tag{23}$$

Let $\varepsilon \in (0, 1/(2\kappa))$. From the assumption (12), for large $k$,

$$\|E_k s_k\| \leq \varepsilon\|s_k\|. \tag{24}$$

Using (18), (21) and (24) in (23), we obtain

$$\|e_{k+1}\| \leq \kappa\ell\eta_k\|e_k\| + 2\kappa\varepsilon\|s_k\| \leq \kappa\ell\eta_k\|e_k\| + 2\kappa\varepsilon\|e_{k+1}\| + 2\kappa\varepsilon\|e_k\|.$$

Hence, if $e_k \neq 0$ for all large $k$, we have

$$\frac{\|e_{k+1}\|}{\|e_k\|} \leq \frac{\kappa\ell\eta_k + 2\kappa\varepsilon}{1 - 2\kappa\varepsilon}.$$

Since $\eta_k \to 0$ is $\varepsilon$ can be arbitrarily small we obtain q-superlinear convergence of $x_k$ to $\bar{x}$ and the proof is complete. □

When $C = \mathbb{R}^n$ we have that $N_K(x) = \{0\}$ for any $x$ and the Theorem 5 follows from Theorem 6. When $f$ is strictly differentiable at $\bar{x}$ we obtain as a corollary the following Dennis-Moré type theorem for inexact quasi-Newton methods applied to a smooth variational inequality.

**Corollary 7.** *Let $\bar{x}$ be a solution of the variational inequality* (7) *and let $f$ be strictly differentiable at $\bar{x}$ with Jacobian $\nabla f(\bar{x})$. Let $\{B_k\}$ be a sequence of $n \times n$ matrices and let $\eta_k \to 0$. Consider a sequence $\{x_k\}$ generated by* (9) *which is convergent to $\bar{x}$. If $x_k \to \bar{x}$ q-superlinearly, then* (11) *is satisfied with $E_k = B_k - \nabla f(\bar{x})$.*

*Conversely, if $(\nabla f(\bar{x}) + N_K)^{-1}(0) = \{0\}$ and condition* (12) *holds with $E_k = B_k - \nabla f(\bar{x})$, then $x_k \to \bar{x}$ q-superlinearly.*

For $C = \mathbb{R}^n$ and $\eta_k$ is the zero sequence, Corollary 7 becomes Theorem 1.

We will now show if the function $f$ is not only Lipschitz continuous around $\bar{x}$ but also *semismooth* at $\bar{x}$, then the particular generalized Jacobian $A_k \in \partial f(\bar{x})$ which satisfies (4) and appears in (11) can be replaced by *any* $A_k$ which belongs to either $\partial f(\bar{x})$ or $\partial f(x_k)$ for all $k$. On the other hand, under strong subregularity of $f + N_C$, if (12) holds with $E_k = B_k - A_k$ for every $A_k \in \partial f(x_k)$ then we have q-superlinear convergence. Clearly, if (12) holds with

$E_k = B_k - A_k$ for any choce of $A_k \in \partial f(\bar{x})$, Theorem 6 yields q-superlinear convergence. Recall that a function $f : \mathbb{R}^n \to \mathbb{R}^m$ is semismooth at $\bar{x}$ when it is Lipschitz continuous around $\bar{x}$, directionally differentiable in any direction, and for every $\varepsilon > 0$ there exists $\delta$ such that for every $x \in I\!B_\delta(\bar{x})$ and for every $A \in \partial f(x)$ one has

$$\|f(x) - f(\bar{x}) - A(x - \bar{x})\| \leq \varepsilon \|x - \bar{x}\|.$$

Our next result is a Dennis-Moré theorem for semismooth functions.

**Theorem 8.** *Let $\bar{x}$ be a solution of the variational inequality* (7) *and let $f$ be semismooth at $\bar{x}$. Let $\{B_k\}$ be a sequence of $n \times n$ matrices and let $\eta_k \to 0$. Consider a sequence $x_k$ generated by* (9) *which is q-superlinearly convergent to $\bar{x}$. Then, for every sequence of matrices $\{A_k\}$ such that either $A_k \in \partial f(x_k)$ for all $k$ or $A_k \in \partial f(\bar{x})$ for all $k$, condition* (11) *holds with $E_k = B_k - A_k$.*

*Conversely, let the mapping $f + N_C$ is strongly subregular at $\bar{x}$ for* 0 *and consider a sequence $\{x_k\}$ generated by* (9) *which is convergent to $\bar{x}$. If condition* (12) *holds with $E_k = B_k - A_k$ for every sequence of matrices $\{A_k\}$ such that $A_k \in \partial f(x_k)$ for all $k$, then $x_k \to \bar{x}$ q-superlinearly.*

**Proof.** Let $x_k \to \bar{x}$ *q*-superlinearly and let $\varepsilon > 0$. Consider first a sequence $\{A_k\}$ of matrices $A_k \in \partial f(x_k)$ for all $k$. Repeat the proof of Theorem 6 until (17) where we write instead

$$f(\bar{x}) - f(x_k) - A_k s_k + y_k \in E_k s_k + N_K(e_{k+1}). \tag{25}$$

Since the generalized Jacobian $\partial f$ is upper semi-continuous and compact-valued, there exists a constant $\lambda$ such that

$$\|A_k\| \leq \lambda \quad \text{for all } k. \tag{26}$$

The semismoothness of $f$ yields

$$\|f(\bar{x}) - f(x_k) - A_k e_k\| \leq \varepsilon \|e_k\|. \tag{27}$$

Then, from (19), (26) and (27) we obtain

$$\begin{aligned}
&\|f(\bar{x}) - f(x_k) - A_k s_k + y_k\| \\
&\quad \leq \|y_k\| + \|f(x_k) - f(\bar{x}) - A_k e_k\| + \|A_k\|\|e_{k+1}\| \\
&\quad \leq 2\eta_k \ell \|s_k\| + \varepsilon \|e_k\| + \lambda \|e_{k+1}\| \\
&\quad \leq 2\eta_k \ell \|s_k\| + (\lambda + 2)\varepsilon \|s_k\|.
\end{aligned}$$

The inclusion (25) then implies

$$d(0, E_k s_k + N_K(e_k)) \leq 2\eta_k \ell \|s_k\| + (\lambda + 2)\varepsilon \|s_k\|.$$

Since $\eta_k \to 0$ is $\varepsilon$ can be arbitrarily small, we obtain (11).

Consider next a sequence $\{A_k\}$ with $A_k \in \partial f(\bar{x})$ for all $k$. From the upper semicontinuity of $\partial f$, for all $k$ sufficiently large there exists $\bar{A}_k \in \partial f(x_k)$ such that

$$\|A_k - \bar{A}_k\| < \varepsilon. \tag{28}$$

Adding and subtracting $\bar{A}_k$ in (25) we write

$$f(\bar{x}) - f(x_k) - \bar{A}_k s_k - (A_k - \bar{A}_k)s_k + y_k \in E_k s_k + N_K(e_{k+1}). \tag{29}$$

Repeating the argument used in the preceding case, we get

$$\begin{aligned}
&\|f(\bar{x}) - f(x_k) - \bar{A}_k s_k - (A_k - \bar{A}_k)s_k + y_k\| \\
&\quad \le \|y_k\| + \|f(x_k) - f(\bar{x}) - \bar{A}_k e_k\| + \|A_k - \bar{A}_k\|\|s_k\| + \|\bar{A}_k\|\|e_{k+1}\| \\
&\quad \le 2\eta_k \ell \|s_k\| + (\lambda + 3)\varepsilon\|s_k\|.
\end{aligned}$$

This again gives us (11).

Now, suppose that the mapping $f + N_C$ is strongly subregular at the solution $\bar{x}$ for 0 with constant $\kappa$ and the sequence $x_k$ is generated by (9) and convergent to $\bar{x}$. Let (12) hold for every sequence $\{A_k\}$ with $A_k \in \partial f(x_k)$. As in the proof of Theorem 6, there exists $y_k$ satisfying (15) and (16) such that

$$y_k - f(x_k) - B_k s_k + f(x_{k+1}) \in f(x_{k+1}) + N_C(x_{k+1}). \tag{30}$$

Let $\varepsilon \in (0, 1/(3\kappa))$ and let $\tilde{A}_k \in \partial f(\bar{x})$ be a sequence of matrices satisfying (4) with this $\varepsilon$ for all $k$ sufficiently large. Then, from the upper semicontinuity of $\partial f$ there exists $A_k \in \partial f(x_k)$ such that, for large $k$,

$$\|A_k - \tilde{A}_k\| \le \varepsilon. \tag{31}$$

By assumption, (12) holds with $E_k = B_k - A_k$. Then, for all $k$ large enough we have

$$\|E_k s_k\| \le \varepsilon\|s_k\| \tag{32}$$

and, from (4),

$$\| - f(x_k) - \tilde{A}_k s_k + f(x_{k+1})\| \le \varepsilon\|s_k\|. \tag{33}$$

The strong subregularity of $f + N_C$ yields the inequality (22) for large $k$. Rewriting (30) as

$$y_k - E_k s_k - (A_k - \tilde{A}_k)s_k - f(x_k) - \tilde{A}_k s_k + f(x_{k+1}) \in f(x_{k+1}) + N_C(x_{k+1}).$$

and using (16), (22), (31), (32) and (33) we obtain

$$\begin{aligned}
\frac{1}{\kappa}\|e_{k+1}\| &\le \|y_k\| + \|E_k s_k\| + \|A_k - \tilde{A}_k\|\|s_k\| + \| - f(x_k) - \tilde{A}_k s_k + f(x_{k+1})\| \\
&\le \eta_k \ell \|e_k\| + \varepsilon\|s_k\| + \varepsilon\|s_k\| + \varepsilon\|s_k\| \\
&\le \eta_k \ell \|e_k\| + 3\varepsilon\|e_{k+1}\| + 3\varepsilon\|e_k\|.
\end{aligned}$$

Hence, if $e_k \neq 0$ for all $k$,

$$\frac{\|e_{k+1}\|}{\|e_k\|} \le \frac{\kappa\eta_k\ell + 3\kappa\varepsilon}{1 - 3\kappa\varepsilon}.$$

Since $\eta_k \to 0$ and $\varepsilon$ can be arbitrarily small, we obtain q-superlinear convergence. □

Condition (12) obviously implies (11) since the normal cone always contains the origin. It is an open question how far conditions (11) and (12) are from each other.

As a corollary we obtain the following Dennis-Moré theorem for semismooth equations.

**Corollary 9** (semismooth Dennis-Moré)**.** *Suppose that $f$ is semismooth at $\bar{x}$, a zero of $f$ and satisfies* (6)*, that is, $f$ is strongly subregular at $\bar{x}$. Let $\{B_k\}$ be a sequence of matrices and consider a sequence $\{x_k\}$ generated by* (1) *and converget to $\bar{x}$. Then $x_k \to \bar{x}$ q-superlinearly if and only if for every sequence $\{A_k\}$ of matrices such that either $A_k \in \partial f(x_k)$ for all $k$ or $A_k \in \partial f(\bar{x})$ for all $k$ one has*

$$\lim_{k\to\infty} \frac{\|(B_k - A_k)s_k\|}{\|s_k\|} = 0.$$

From Theorem 8 we also obtain the following result.

**Corollary 10.** *Consider the variational inequality* (7) *with a solution $\bar{x}$ at which the function $f$ is semismooth and the mapping $f + N_C$ is strongly subregular at $\bar{x}$ for* 0*. Consider a sequence $\{x_k\}$ generated by the inexact semismooth Newton method* (9) *with $B_k = A_k$ for any matrix $A_k \in \partial f(x_k)$, $k = 0, 1, 2, \ldots$ If $\{x_k\}$ is convergent to $\bar{x}$, then it is convergent q-superlinearly.*

When $C = \mathbb{R}^n$, then the strong subregularity assumption reduces to condition (6) which, combined with the semismoothness actually implies that the generalized Jacobian $\partial f(\bar{x})$ contains nonsingular matrices only. Then we recover the standard setting for proving convergence of semismooth Newton methods broadly exhibited in the books [8], [9] and [11].

At the end of the paper we present an erratum to our previous paper [3] where we considered solving equations involving functions acting between Banach spaces $X$ and $Y$ that have the following property around a point $\bar{x} \in \operatorname{int}\operatorname{dom} f$: there exist a neighborhood $U$ of $\bar{x}$ and a set-valued mapping $\mathcal{A} : U \rightrightarrows \mathcal{L}(X,Y)$, the space of linear and bounded mappings from $X$ to $Y$, such that

$$\sup_{A\in\mathcal{A}(x)} \|f(x) - f(\bar{x}) - A(x-\bar{x})\| = o(\|x-\bar{x}\|) \quad \text{as} \quad x \to \bar{x}.$$

This class of functions was apparently introduced by B. Kummer who called them Newton mappings, a name which later propagated in the literature as Newton differential functions. Xu [12] defined this class as functions having a point-based set-valued approximation, while in his recent book Penot [10] used the name slantly differentiable functions. In [3] the author named this kind of differentiability after Kummer, with the intention to give credit to the individual who introduced it. As it turns out, however, *every function is Kummer/Newton/point-based/slant differentiable.* This simple fact is explicitely shown in the recent book of Penot [10, Lemma 2.64], but pehaps well known much earlier since a finite-dimensional version of it was given in [12] and credited there to a referee of that paper.

In author's previous paper we presented a Dennis-Moré theorem, [3, Theorem 2], for equations involving the class of nonsmooth functions in Banach spaces described above. Unfortunately, there is a gap in the proof of this theorem and it an open problem whether its statement is correct. Below we prove a corrected version of this result involving an additional assumption.

**Theorem 11.** ([3, Theorem 2] corrected) *Let $X$ and $Y$ be Banach spaces. Consider a function $f : X \to Y$ with a zero $\bar{x}$. Let for some starting point $x_0$ the sequence $\{x_k\}$ be generated by the method* (1) *and convergent to $\bar{x}$ for a sequence $\{B_k\}$ of linear and bounded mappings $B_k : X \to Y$. Let $\{A_k\}$ be a bounded sequence of mappings $A_k \in \mathcal{L}(X,Y)$ with the property that for every $\varepsilon > 0$ there exists $\bar{k}$ such that*

$$\|f(x_k) - f(\bar{x}) - A_k(x_k - \bar{x})\| \le \varepsilon\|x_k - \bar{x}\| \quad \textit{for all} \quad k \ge \bar{k}. \tag{34}$$

*Let $E_k = B_k - A_k$. Then the following implications hold:*

*(i) If $x_k \to \bar{x}$ q-superlinearly then*

$$\lim_{k\to\infty} \frac{\|E_k s_k\|}{\|s_k\|} = 0. \tag{35}$$

*(ii) Conversely, if $f$ satisfies condition* (6)*, that is, $f$ is strongly subregular at $\bar{x}$, and the sequence $\{A_k\}$ is chosen such that*

$$\lim_{k\to\infty} \frac{\|(A_{k+1} - A_k)e_{k+1}\|}{\|e_k\|} = 0, \tag{36}$$

*then condition* (35) *implies that $x_k \to \bar{x}$ q-superlinearly.*

**Proof.** Let $x_k \to \bar{x}$ $q$-superlinearly and choose $\varepsilon > 0$. From the boundedness of the sequence $\{A_k\}$, say by a constant $L$, and condition (14), we have

$$\|A_k e_{k+1}\| \leq 2\varepsilon L \|s_k\| \quad \text{for all } k \text{ sufficiently large.}$$

From (34),

$$\|f(x_k) - A_k e_k\| \leq \varepsilon \|e_k\| \leq 2\varepsilon \|s_k\| \quad \text{for all } k \text{ sufficiently large.}$$

Using these estimates in the equality

$$E_k s_k + A_k e_{k+1} + f(x_k) - A_k e_k = 0 \tag{37}$$

we obtain

$$\|E_k s_k\| \leq 2\varepsilon(L + \varepsilon + 1)\|s_k\|$$

for all $k$ sufficiently large. This yields (35).

Conversely, let $f$ be strongly subregular at $\bar{x}$ with a constant $\kappa$ and let (36) hold. Choose $\varepsilon \in (0, 1/(2\kappa))$. Then, for sufficiently large $k$,

$$\|E_k s_k\| \leq \varepsilon \|s_k\|. \tag{38}$$

and

$$\|(A_{k+1} - A_k)e_{k+1}\| \leq \varepsilon \|e_k\|. \tag{39}$$

Taking into account the strong subregularity of $f$ and (34), we have

$$\|A_{k+1}e_{k+1}\| \geq \|f(x_{k+1})\| - \|f(x_{k+1} - f(\bar{x}) - A_{k+1}e_{k+1}\| \geq (1/\kappa - \varepsilon)\|e_{k+1}\|. \tag{40}$$

On the other hand, since $E_k s_k + A_k e_{k+1} f(x_k) - A_k e_k = 0$, using (34), (38) and (39), we obtain

$$\begin{aligned}
\|A_{k+1}e_{k+1}\| &\leq \|(A_{k+1} - A_k)e_{k+1}\| + \|A_k e_{k+1}\| \\
&\leq \varepsilon\|e_{k+1}\| + \|E_k s_k\| + \|f(x_k) - f(\bar{x}) - A_k e_k\| \\
&\leq \varepsilon\|e_{k+1}\| + \varepsilon\|s_k\| + \varepsilon\|e_k\| \\
&\leq 2\varepsilon\|e_{k+1}\| + 2\varepsilon\|e_k\|.
\end{aligned}$$

Combining (40) with the last inequality, we get

$$\|e_{k+1}\| \leq \frac{2\varepsilon}{1 - 2\kappa\varepsilon}\|e_k\|$$

Since $\varepsilon$ is arbitrarily small, this yields q-superliner convergence. ◻

It an open question to identify specific quasi-Newton methods for which the conditions for q-superliner convergence displayed in the theorems above are satisfied.

## References


[1] J. E. Dennis, Jr., J. J. Moré, A characterization of superlinear convergence and its application to quasi-Newton methods, Math. Comp. **28** (1974), 549–560.

[2] R. S. Dembo, S. C. Eisenstat, T. Steihaug, Inexact Newton methods. SIAM J. Numer. Anal. **19** (1982), 400–408.

[3] A. L. Dontchev, Generalizations of the Dennis-Moré theorem, *SIAM J. Optim.* **22** (2012) 821–830.

[4] A. L. Dontchev, R. T. Rockafellar, Convergence of inexact Newton methods for generalized equations, Mathematical Programming B, accepted.

[5] A. L. Dontchev, R. T. Rockafellar, *Implicit functions and solution mappings,* Springer, Dordrecht 2009.

[6] M. Fabian, Concerning interior mapping theorem, Comment. Math. Univ. Carolin. **20** (1979) 345–356.

[7] C. T. Kelley, Solving nonlinear equations with Newton's method, Fundamentals of Algorithms, SIAM, Philadelphia, PA, 2003.

[8] K. Ito, K. Kunisch, *Lagrange multiplier approach to variational problems and applications,* SIAM, Philadelphia, PA, 2008.

[9] D. Klatte, B. Kummer, *Nonsmooth equations in optimization. Regularity, calculus, methods and applications,* Kluwer, Dordrecht, 2002.

[10] J.- P. Penot, *Calculus without derivatives,* Springer, Dordrecht 2013.

[11] M. Ulbrich, *Semismooth Newton methods for variational inequalities and constrained optimization problems in function spaces*, SIAM, Philadelphia, PA, 2011.

[12] H. Xu, *Set-valued approximations and Newton's method,* Math. Programming A **84** (1999) 401–420.