ON THE CONVERGENCE OF THE NEWTON ITERATION

K. VESELIĆ

ABSTRACT. We give a new class of convergence criteria for the classical Newton method for the equation $f(x) = 0$. We prove the global convergence, if the function f is the gradient of a uniformly convex function, while the convergence in the general case is still open. Our convergence condition consists of only one bound with a constant, which is shown to be sharp. We apply our result to a generic variational problem as well as to uniformly diagonally dominant Hessians. The convergence result is extended to the cases where the derivative is defined merely as a distribution.

1. INTRODUCTION AND PRELIMINARIES

In this paper we consider the classical Newton process

$$x_{k+1} = x_k - Df(x_k)^{-1} f(x_k) \tag{1}$$

to solve the equation

$$f(x) = 0,$$

where

$$f : \Omega \to X$$

and Ω is an open connected subset of a real Hilbert space X with the scalar product $x^T y$ and the norm $\|\cdot\|$. Here Df is the standard notation for the Fréchet derivative of f.[1]

Many existing criteria (see e.g. [3], [6], [1], [2] and the references cited there) use some sort of continuity of Df to prove the convergence in some neighbourhood of the zero. Others, like [8], [7] allow f to be differentiable in some weak sense. Our choice is to compare Hessians in their natural quadratic form ordering. This leads us first to the special case in which f is the gradient of a twice continuously differentiable function V with $D^2V(x)$ uniformly positive definite[2]; we require

$$K = \sup_{\xi \neq 0,\, x,y \in \Omega} \frac{\xi^T D^2 V(x) \xi}{\xi^T D^2 V(y) \xi} < 2\,. \tag{2}$$

This simple condition, which seems to be new, handsomely expresses the requirement that "$D^2V(x)$ should not vary too much with x" which will prevent the Newton iteration from 'overshooting'. Indeed, the possible values of K lie between 1 and 2: the case $K < 1$ is obviously impossible and for $K = 1$ the function V is trivially quadratic, while for any $K > 2$ a simple non-convergent case will be constructed.

[1] We take this notation for the scalar product, to keep analogy with the case of finite matrices. In this case Df will be understood in its natural matrix representation.

[2] By uniform positive definiteness we mean $\inf_{\xi \neq 0} \xi^T D^2 V(x) \xi / \xi^T \xi > 0$ (for finite matrices one usually says just 'positive definite'). In fact, if valid for some x the uniform positive definiteness extends to all $x \in \Omega$ by virtue of (2).

The practical value of the classical Newton iteration is known to be limited and in practice it is often conveniently dampened in order to insure global convergence (cf. e.g. [5]). Nevertheless, better understanding of the mechanism and the geometry of the classical method may be useful in such environments as well.

Our condition is readily extended to the cases in which $Df = D^2V$ exists only in the distributional sense. In this case the space X has to be finite dimensional. The generalized derivatives in [8], [7] are less general than ours so that they still allow the space X to be infinite dimensional. There is little doubt that our condition will work within such frameworks as well. We chose the distributional setting because of its generality, it proved to be in nice harmony with convexity arguments in this context. Also, as could be expected, the simplest non-convergent counterexamples have $Df = D^2V$ which is only piecewise continuous.

Our result seems to be new even in the one-dimensional case. Indeed, for $\Omega = R$ the condition (2) can be written as

$$0 < \sup f' < 2 \inf f'.$$

As an application of our new bound, we show the global convergence in the case of the uniform diagonal dominance of the Hessian matrix:

$$D^2V(x) = A + N(x),$$

where A is a fixed positive definite matrix and

$$|\xi^T N(x)\xi| \le K_2 \xi^T A \xi, \quad K_2 < 1/3.$$

Another application is the global convergence for a generic nonlinear variational problem

$$V(u) = \int_0^1 \left(\frac{1}{2}u'(x)^2 + f(u(x))\right) dx = \min$$

under the condition $\sup f'' - 2\inf f'' < \pi^2 < \inf f''$. In the last section we discuss some possibilities of generalizing our results to the non-convex case.

2. Newton minimizing of a convex function V

Let V be a twice continuously differentiable real-valued function on a Hilbert space X. The classical Newton method (1) reads

$$x_{j+1} = x_j - D^2V(x_j)^{-1} DV(x_j), \tag{3}$$

where $f = DV$ and $Df = D^2V$. The latter is understood as a bounded symmetric uniformly positive definite operator in X.

An equivalent formulation of (2) is

$$K = \sup_{\xi \ne 0,\, x,y \in \Omega} \frac{\xi^T D^2V(x)^{-1}\xi}{\xi^T D^2V(y)^{-1}\xi} < 2\,. \tag{4}$$

Indeed, starting from (2) and using $\|A^*A\| = \|AA^*\|$ we have

$$K = \sup_{x,y} \|D^2V(y)^{-1/2} D^2V(x) D^2V(y)^{-1/2}\| = \sup_{x,y} \|D^2V(x)^{1/2} D^2V(y)^{-1} D^2V(x)^{1/2}\|$$

and that is K from (4). Another way to interpret this condition uses (in the finite dimensional case) known extremal properties of the eigenvalues: the zeros of the secular equation

$$\det(D^2V(x) - \lambda D^2V(y)) = 0$$

lie in the interval $[1/K, K]$ as x, y run over Ω.

In what follows we will repeatedly use the second order Taylor formula connecting the points x_0 and x

$$V(x) = V(x_0) + (x - x_0)^T DV(x_0) + \frac{1}{2}(x - x_0)^T D^2V(\eta)(x - x_0), \tag{5}$$

where η lies between x_0 and x. We will use the same generic symbol η in any Taylor formula in the sequel.

Theorem 2.1. *Let V be twice continuously differentiable on an open convex set $\Omega \subseteq X$ with D^2V uniformly positive definite and satisfying (2). Let $x_1 \in \Omega$ with $DV(x_1) \neq 0$,*

$$\mathcal{V} = \{x \in \Omega : V(x) \leq V(x_1)\} \tag{6}$$

and suppose

$$\partial\mathcal{V} \cap \partial\Omega = \emptyset, \tag{7}$$

where $\partial\Omega$, $\partial\mathcal{V}$ is the boundary of Ω, $\mathcal{V}$, respectively. Then all Newton iterates stay in $\mathcal{V}$ and they converge to the unique minimum point x^* of V in Ω. The convergence is monotone i.e. $V(x_{j+1}) < V(x_j)$ unless $x_{j+1} = x_j$.

Proof. The convexity of $\mathcal{V}$ follows from the obvious convexity of the function V on Ω. For any $x, y \in \Omega$ (2) implies

$$V(x) \geq V(y) + (x - y)^T DV(y) + \frac{1}{2K}(x - y)^T D^2V(y)(x - y),$$

so the uniform positive definiteness of $D^2V(y)$ implies that $\mathcal{V}$ is bounded and that V is bounded from below on Ω.

From all that it follows that V possesses a unique minimum point x^* in Ω (in fact, $x^* \in \mathcal{V}$); this is the unique solution of the equation $DV(x) = 0$ in Ω.

Let now $x_j \in \Omega$, $DV(x_j) \neq 0$. Then $x_{j+1} - x_j \neq 0$ and we set

$$x(\alpha) = \alpha x_{j+1} + (1 - \alpha)x_j. \tag{8}$$

Let $x(\alpha) \in \Omega$. Using the Taylor formula (5) between x_j and $x(\alpha)$, and the formula (3) we obtain

$$\begin{aligned} V(x(\alpha)) - V(x_j) &= \alpha(x_{j+1} - x_j)^T DV(x_j) + \frac{\alpha^2}{2}(x_{j+1} - x_j)^T D^2V(\eta)(x_{j+1} - x_j) \\ &= DV(x_j)^T \left(-\alpha D^2V(x_j)^{-1} + \frac{\alpha^2}{2} D^2V(x_j)^{-1} D^2V(\eta) D^2V(x_j)^{-1}\right) DV(x_j). \end{aligned}$$

Now we use first (2) with $x = \eta$, $y = x_j$ and then (4) with $x = x_1$, $y = x_j$ to obtain

$$V(x(\alpha)) - V(x_j) \leq -\alpha\left(\frac{1}{K} - \frac{\alpha}{2}\right) DV(x_j)^T D^2V(x_1)^{-1} DV(x_j) < 0, \quad \text{for } 0 < \alpha < 2/K. \tag{9}$$

We will now prove that all x_j lie in $\mathcal{V}$. We proceed by induction. Obviously, $x_1 \in \mathcal{V}$. Suppose $x_j \in \mathcal{V}$. The already established properties of the sets Ω and $\mathcal{V}$, together with (7) and (9) imply

$$\mathcal{V} \cap \{x(\alpha) : \alpha \in R\} = \{x(\alpha) : \beta_1 \le \alpha \le \beta_2\}, \tag{10}$$

$$\Omega \cap \{x(\alpha) : \alpha \in R\} = \{x(\alpha) : \alpha_1 < \alpha < \alpha_2\},$$

with

$$\alpha_1 < \beta_1 \le 0 < \beta_2 < \alpha_2.$$

Now, $x_{j+1} = x(1) \notin \mathcal{V}$ would imply $\beta_2 < 1 < 2/K$, whereas (9) would imply

$$x(\alpha) \in \mathcal{V}, \quad \alpha \in [\beta_2, \min\{2/K, \alpha_2\})$$

— a contradiction to (10). Thus, all x_j lie in $\mathcal{V}$ (and therefore in Ω). By setting $\alpha = 1$ in (9) we obtain $V(x_{j+1}) < V(x_j)$ and

$$DV(x_j) \to 0.$$

We now prove $x_j \to x^*$. Note that (i) the equation $DV(x) = y$ has at most one solution in Ω and (ii) by the inverse function theorem $DV(\cdot)$ is a diffeomorphism of some neighbourhood of x^* onto an open ball containing zero. Now $DV(x_j) \to 0$ implies $x_j \to x^*$. Q.E.D.

Remark 2.1. The condition (7) naturally insures that the Newton iterates stay within Ω. If $\Omega = X$ it is automatically fulfilled and the proof is simpler still.

Example 2.1. Take

$$V(u) = \int_0^1 \left(\frac{1}{2}u'(x)^2 + f(u(x))\right) dx, \quad u(0) = u_0, \quad u(1) = u_1$$

where $f : R \to R$ is twice continuously differentiable. We take X as the first order Sobolev space with the scalar product $u^T v = \int_0^1 u'(x)v'(x)dx$ with u, v vanishing at both ends of the interval. The second derivative of V is given as a sesquilinear form by

$$u^T D^2V(w)v = \int_0^1 (u'(x)v'(x) + f''(w(x))u(x)v(x))\, dx.$$

Our theorem above immediately applies, provided that

$$\inf f'' > \pi^2, \quad \sup f'' - 2\inf f'' < \pi^2.$$

Here we have used the known identity

$$\pi^2 = \min_u \frac{\int_0^1 u'^2 dx}{\int_0^1 u^2 dx}.$$

We see that the quadratic form setting of Theorem 2.1 allows an easy checking of the required conditions without representing the operator $D^2V(x)$ itself, which would involve integral operators.

The fulfillment of the condition (7) may be plausible, but not directly verifiable. In the case of

$$\Omega = \{x : (x - x^*)^T D^2 V(x^*)(x - x^*) < r\}$$

(this is a "$D^2V(x^*)$ ball") the mentioned condition is implied by the inequality

$$V(x_1) < V(x^*) + \frac{r}{2K}. \tag{11}$$

Indeed, take any $y \in \Omega$ with $V(y) = V(x_1)$. Then using the Taylor formula between x^* and y we have

$$(y - x^*)^T D^2 V(x^*)(y - x^*) \le K(y - x^*)^T D^2 V(\eta)(y - x^*) =$$
$$= 2K(V(y) - V(x^*)) = 2K(V(x_1) - V(x^*)) < r.$$

Since $\mathcal{V}$ is convex the estimate above remains valid, if y is replaced by any vector from $\mathcal{V}$. That is, we have $\mathcal{V} \subseteq \Omega$. Since $\mathcal{V}$ is closed and Ω open (7) follows (note that the $D^2(V)$-norm and the standard norm are topologically equivalent).

Another generalisation of the results obtained above is the case where the second derivative of the function V exists only as a distribution (in this case we have to assume the finite dimensionality of X). Assume that $V : \Omega \to R$ is continuously differentiable and let its distributional second derivative satisfy

$$D^2 V \in L^\infty(\Omega) \tag{12}$$

and

$$L\|\xi\|^2 \le \operatorname{ess\,inf} \xi^T D^2 V(\cdot)\xi, \quad \operatorname{ess\,sup} \xi^T D^2 V(\cdot)\xi \le K \operatorname{ess\,inf} \xi^T D^2 V(\cdot)\xi \tag{13}$$

for any $\xi \in R^n$ and $L > 0$, $\quad 1 \le K < 2$.

If the function V is twice continuously differentiable this condition is equivalent to the condition (2) together with the uniform positive definiteness of D^2V.

Let $\varepsilon : R^n \mapsto R$ be a mollifier i.e. an infinitely smooth non-negative function, vanishing outside of the ball $\{x : \|x\| < r_\varepsilon\}$ and satisfying

$$\int \varepsilon(x)dx = 1.$$

A sequence of mollifiers ε is called a δ-sequence (denoted by $\varepsilon \to \delta$), if the radii of their supporting balls tend to zero. Set

$$V_\varepsilon(x) = \int V(y)\varepsilon(x - y)dy. \tag{14}$$

Then

$$DV_\varepsilon(x) = \int DV(y)\varepsilon(x - y)dy, \quad D^2V_\varepsilon(x) = \int D^2V(y)\varepsilon(x - y)dy. \tag{15}$$

It is immediately seen that under our assumptions with any δ-sequence of ε's the function V_ε converges to V uniformly over compacts and similarly for DV. For D^2V there need not be convergence but the set $\mathcal{S}(x)$ of all accumulation points of $D^2V_\varepsilon(x)$ over all δ-sequences of ε's is not void. Moreover, this set is convex. This is readily seen from the fact that all mollifiers obviously make a convex set and from the property

$$D^2V_{\sigma\varepsilon+\tau\eta} = \sigma D^2V_\varepsilon + \tau D^2V_\eta,$$

where ε, η are any mollifiers. We define the Newton process as

$$x_{k+1} = x_k - S_k^{-1} DV(x_k), \quad k = 1, 2, \ldots \tag{16}$$

with $S_k \in \mathcal{S}(x_k)$. Note that this process is in general not uniquely determined by the start value x_1.

Theorem 2.2. *Let V, $\mathcal{S}(x)$, x_k satisfy (12), (13), (16) and let $\Omega = R^n$. Then*
- (i) *V possesses a unique minimum point x^* and*
- (ii) *the process (16) converges to x^* monotonously i.e.*

$$V(x_{k+1}) < V(x_k) \quad \textit{whenever } x_{k+1} \neq x_k.$$

Proof. We first prove that V_ε satisfies (13). Indeed,

$$\xi^T D^2 V_\varepsilon(x)\xi = \int \xi^T D^2 V(y)\xi\varepsilon(x-y)dy \leq$$
$$\leq \operatorname{ess\,sup} \xi^T D^2 V(\cdot)\xi \leq K \operatorname{ess\,inf} \xi^T D^2 V(\cdot)\xi \leq$$
$$\leq K \int \xi^T D^2 V(z)\xi\varepsilon(y-z)dz = K\xi^T D^2 V_\varepsilon(y)\xi$$

for any $\xi, x, y \in R^n$. Also

$$\xi^T D^2 V_\varepsilon(x)\xi \geq \operatorname{ess\,inf} \xi^T D^2 V(\cdot)\xi \geq L\|\xi\|^2.$$

By continuity we have

$$LI \leq A \leq KB, \quad A, B \in \mathcal{S},$$

where $\mathcal{S} = \cap_x \mathcal{S}(x)$ and $\leq$ is the quadratic-form ordering of symmetric matrices.

Next we prove the existence and the uniqueness of the minimum point x^*. The proof is standard and we give it mainly for the sake of completeness. For any $x, y \in R^n$, $\varepsilon > 0$ the Taylor formula gives

$$V_\varepsilon(x) = V_\varepsilon(y) + (x-y)^T DV_\varepsilon(y) + \frac{1}{2}(x-y)^T D^2 V_\varepsilon(\eta)(x-y) \geq$$
$$\geq V_\varepsilon(y) + (x-y)^T DV_\varepsilon(y) + \frac{L}{2}\|x-y\|^2.$$

Taking $\varepsilon \to \delta$ we obtain

$$V(x) \geq V(y) + (x-y)^T DV(y) + \frac{L}{2}\|x-y\|^2.$$

Thus, $V(x)$ is bounded from below and grows to infinity as $\|x\| \to \infty$. Hence there exists a minimum point x^* with $DV(x^*) = 0$. If y^* is another such point then using (2) for V_ε we have

$$V_\varepsilon(x^*) = V_\varepsilon(y^*) + (x^*-y^*)^T DV_\varepsilon(y^*) + \frac{1}{2}(x^*-y^*)^T D^2 V_\varepsilon(\eta)(x^*-y^*) \geq$$
$$\geq V_\varepsilon(y^*) + (x^*-y^*)^T DV_\varepsilon(y^*) + \frac{L}{2}\|x^*-y^*\|^2$$

and, taking again $\varepsilon \to \delta$

$$0 = V(x^*) - V(y^*) \geq \frac{L}{2}\|x^*-y^*\|^2$$

and

$$x^* = y^*.$$

Thus, the minimum position exists and is unique.

Consider now any sequence x_k satisfying (16). For any k take any mollifier ε. Then

$$V_\varepsilon(x_{j+1}) - V_\varepsilon(x_j) = (x_{k+1} - x_k)^T DV_\varepsilon(x_k) + \frac{1}{2}(x_{k+1} - x_k)^T D^2 V_\varepsilon(\eta)(x_{k+1} - x_k) \leq$$

$$\leq (x_{k+1} - x_k)^T DV_\varepsilon(x_k) + \frac{K}{2}(x_{k+1} - x_k)^T D^2 V_\varepsilon(x_k)(x_{k+1} - x_k).$$

Taking a sequence $\varepsilon \to \delta$ such that $D^2 V_\varepsilon(x_k) \to S_k \in \mathcal{S}(x_k)$ and using (16) gives

$$V(x_{k+1}) \leq V(x_k) - \left(1 - \frac{K}{2}\right)(x_{k+1} - x_k)^T S_k (x_{k+1} - x_k) =$$

$$= V(x_k) - \left(1 - \frac{K}{2}\right) DV(x_k)^T S_k^{-1} DV(x_k) \leq$$

$$\leq V(x_k) - \left(\frac{1}{K} - \frac{1}{2}\right) DV(x_k)^T S_1^{-1} DV(x_k).$$

Thus, $V(x_{k+1}) < V(x_k)$ except for $x_{k+1} = x_k$ and therefore $\lim_{k\to\infty} V(x_k) = V^*$ exists and, by the uniform positive definiteness of S_1,

$$\lim_{k\to\infty} DV(x_k) = 0.$$

The sequence x_k is bounded. Indeed, the Taylor formula connecting x_k and x^* gives

$$V_\varepsilon(x_k) \geq V_\varepsilon(x^*) + (x_k - x^*) DV_\varepsilon(x^*) + \frac{L}{2}\|x_k - x^*\|^2.$$

Now, let $\varepsilon \to \delta$ and use $DV(x^*) = 0$:

$$V(x_k) \geq V(x^*) + \frac{L}{2}\|x_k - x^*\|^2.$$

Altogether, x_k converges to x^*. Q.E.D.

A corresponding local version as in Theorem 2.1 could be proved as well, we omit the details.

Example 2.2. (Diagonal dominance) Let $\Omega = R^n$ and

$$D^2 V(x) = A + N(x)$$

with $N(\cdot)$ measurable, A positive definite and

$$|\xi^T N(x) \xi| \leq K_2 \xi^T A \xi, \quad K_2 < 1/3$$

for all ξ and almost all x. Then

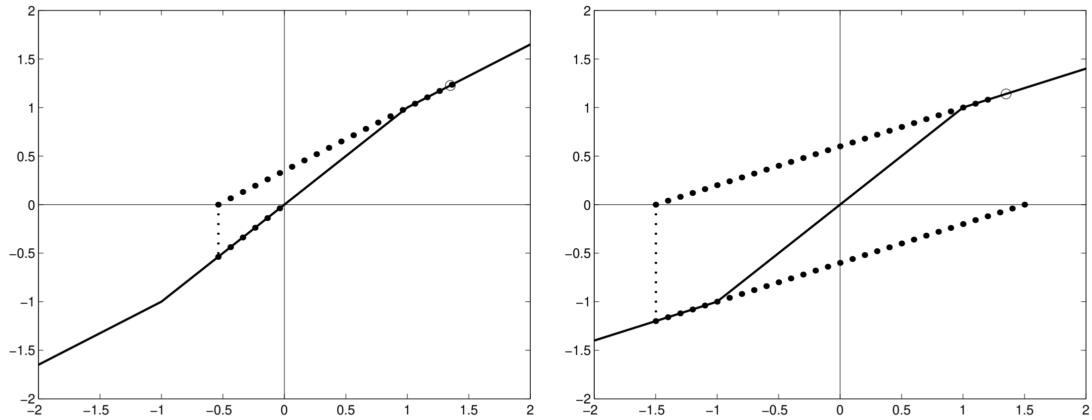
$$\xi^T D^2 V(x) \xi \leq (1 + K_2) \xi^T A \xi, \quad \xi^T D^2 V(x) \xi \geq (1 - K_2) \xi^T A \xi.$$

Thus, for any x, y,

$$\xi^T D^2 V(x) \xi \leq (1 + K_2) \xi^T A \xi \leq \frac{1 + K_2}{1 - K_2} \xi^T D^2 V(y) \xi$$

and
$$\frac{1+K_2}{1-K_2} < 2.$$
Of course, we may make a linear transformation of coordinates
$$x' = A^{1/2}x$$
such that A becomes the identity matrix; in this representation our condition is insured by the "1/3-diagonal dominance" of D^2V.

The following examples show that the bound $K < 2$ is sharp in $X = R^n$. The first example deals with $X = R$ and is illuminating in its simplicity; it serves as a basis for the construction of examples in arbitrary dimensions. We bring these examples in some detail, because they throw additional light on the character of our main convergence condition.



(a) (b)

FIGURE 1. Convergent process (a), Divergent process (b)

Example 2.3. Take $0 < \alpha < 1$ and set
$$f(x) = DV(x) = \begin{cases} \alpha x + 1 - \alpha, & x \geq 1 \\ x, & |x| \leq 1 \\ \alpha x - 1 + \alpha, & x \leq -1 \end{cases}$$
Then
$$f'(x) = \begin{cases} \alpha, & |x| > 1 \\ 1, & |x| < 1 \end{cases}$$
Obviously, (12), (13) holds with
$$K = \frac{1}{\alpha}$$
and Theorem 2.2 applies for $\alpha > 1/2$. In this case the process converges in at most 2 steps (Fig. 2 (a)). For $K > 2$ we have $\alpha < 1/2$ and for $|x_1| > 1$ the process becomes periodic without convergence since the equation
$$\alpha x \pm (1 - \alpha) = 0$$

has the zeros

$$x_{\pm} = \mp(\frac{1}{\alpha} - 1)$$

and they are absolutely larger than 1 (Fig. 2 (b)).

Example 2.4. Take the function from Example 2.3 with $1/2 < \alpha < 1$ and for $|x| \geq 2$ replace it by

$$f(x) = \begin{cases} \beta(x-2) + 1 + \alpha, & x \geq 2 \\ \beta(x+2) - 1 - \alpha, & x \leq -2 \end{cases}$$

with $0 < \beta < 1/2$. The Newton process converges when started in $\Omega = (-2, 2)$ and diverges when started with $x \notin \bar{\Omega}$.

The generalisation to the space X of arbitrary dimension is straightforward, but less trivial:

Example 2.5. For the case $\Omega = X = R^n$ we take

$$V(x) = \begin{cases} \|x\|^2/2, & \|x\| < 1 \\ \alpha\|x\|^2/2 + (1-\alpha)(\|x\| - 1/2), & \|x\| \geq 1 \end{cases} \tag{17}$$

with $1/2 < \alpha < 1$. Hence

$$DV(x) = \begin{cases} x, & \|x\| < 1 \\ \gamma(x)x, & \|x\| \geq 1 \end{cases}$$

$$D^2V(x) = \begin{cases} I, & \|x\| < 1 \\ \gamma(x)I - \zeta(x)\frac{xx^T}{\|x\|^2}, & \|x\| > 1 \end{cases}$$

with

$$\gamma(x) = \alpha + \frac{1-\alpha}{\|x\|}, \quad \zeta(x) = \frac{1-\alpha}{\|x\|}.$$

The obtained D^2V is piecewise continuous, so this is simultaneously the second derivative of V in the sense of distributions. The Newton iteration reads

$$x' = x - D^2V(x)^{-1}DV(x) = \begin{cases} 0, & \|x\| < 1 \\ -\frac{1-\alpha}{\alpha}\frac{x}{\|x\|}, & \|x\| > 1 \end{cases}$$

that is, as in the one-dimensional example above, the Newton process diverges for $\alpha < 1/2$. It remains to prove that V satisfies the conditions (12) and (13) with $K = 1/\alpha$. The first one is obvious. We will prove the second one, if we prove that the matrix

$$A(x, y) = D^2V(y) - \alpha D^2V(x)$$

is positive semidefinite for arbitrary $x, y \in R^n$ with $\|x\|, \|y\| \neq 1$.

In the case $\|x\| < 1, \|y\| < 1$ we have $A(x, y) = (1-\alpha)I$ which is obviously positive semidefinite.

In the case $\|x\| < 1, \|y\| > 1$ we have

$$A(x,y) = (\gamma(y) - \alpha)I - \zeta(y)\frac{yy^T}{\|y\|^2}$$

and this is positive semidefinite by

$$\gamma(y) - \alpha - \zeta(y) = 0.$$

In the case $\|x\| > 1, \|y\| < 1$ we have

$$A(x,y) = I - \alpha\gamma(x)I + \alpha\zeta(x)\frac{xx^T}{\|x\|^2}$$

and this is positive semidefinite by

$$1 - \alpha\gamma(x) = 1 - \alpha^2 - \frac{\alpha(1-\alpha)}{\|x\|} \geq 1 - \alpha^2 - \alpha + \alpha^2 = 1 - \alpha > 0.$$

In the case $\|x\| > 1, \|y\| > 1$ we have

$$A(x,y) = (\gamma(y) - \alpha\gamma(x))I - \zeta(y)vv^T + \alpha\zeta(x)uu^T,$$

$$u = \frac{x}{\|x\|}, \quad v = \frac{y}{\|y\|}.$$

If u and v are parallel then the positive semidefiniteness of $A(x,y)$ follows from

$$\gamma(y) - \alpha\gamma(x) = \alpha + \frac{1-\alpha}{\|y\|} - \alpha^2 - \frac{\alpha(1-\alpha)}{\|x\|} \geq$$

$$\alpha - \alpha^2 - \alpha(1-\alpha) = 0$$

and

$$\gamma(y) - \alpha\gamma(x) - \zeta(y) + \alpha\zeta(x) = \alpha - \alpha^2 > 0.$$

If u and v are not parallel then on their orthogonal complement A equals the scalar $\gamma(y) - \alpha\gamma(x)$ which is non-negative (see above). It remains to check the restriction of A to the linear span of u, v or, equivalently, the matrix

$$\begin{bmatrix} u^T \\ v^T \end{bmatrix} A(x,y) \begin{bmatrix} u & v \end{bmatrix} = \begin{bmatrix} \alpha(1-\alpha) + \zeta(y)(1-s^2) & \alpha(1-\alpha)s \\ \alpha(1-\alpha)s & \alpha(1-\alpha) - \alpha\zeta(x)(1-s^2) \end{bmatrix}$$

with $s = u^T v$ and $|s| \leq 1$. This matrix is positive semidefinite being a sum of two obviously positive semidefinite matrices

$$\alpha(1-\alpha)\begin{bmatrix} 1 & s \\ s & 1-(1-s^2)/\|x\| \end{bmatrix} \text{ and } (1-\alpha)\begin{bmatrix} (1-s^2)/\|y\| & 0 \\ 0 & 0 \end{bmatrix}.$$

Thus, we have proved, that V satisfies the conditions of Theorem 2.2 with $K = 1/\alpha < 2$ and that the process diverges, if in (13) the constant K is larger than 2. The generalisation of this example in the sense of Example 2.4 is straightforward. In both cases, V can be immediately mollified to yield a counterexample in the setting of the classical Theorem 2.1 at least in the finite dimensional case.

The example above is valid in the infinite dimensional case as well, only there our convergence theory is not applicable, since D^2V is not continuous and the standard mollifying is impossible because of the lacking of the appropriate integration on infinite dimensional Hilbert spaces. However, a mollifying over the radial variable is possible (our V is spherically symmetric) thus showing that the bound $K<2$ is sharp in the infinite dimensional case as well. We omit the details.

It is desirable to give an intrinsic characterisation of the sets $\mathcal{S}(x)$ since it is out of them that the matrices S_k in (16) can be freely taken (note that in the proof above the mollifier ε can be freely chosen for any k). We will show that in the case of piecewise continuous D^2V the set $\mathcal{S}(x)$ contains any convex combination of all piecewise limits of D^2V at x. The Proposition below is fairly plausible and we bring it mainly for the sake of completeness.

Proposition 2.3. *Let x_0 be such that D^2V has a 'cone limit' V_0 at x_0 i.e.*

$$\operatorname{ess}\sup_{\substack{\|y-x_0\|\le\zeta\\ y-x_0\in\mathcal{C}}}\|D^2V(y)-V_0\|\to 0,\quad \textit{for}\quad \zeta\to 0,$$

where

$$\mathcal{C}=\mathcal{C}(u,\beta)=\{x=\alpha u+x',\ x'\perp u,\ \alpha>\beta\|x'\|\}$$

is a cone defined by the unit vector u and $\beta>0$. Then the set $\mathcal{S}(x_0)$ contains V_0.

Proof. Denote by χ the characteristic function of the cone $\mathcal{C}$ and define the mollifier

$$j(x)=\theta\begin{cases} e^{-\frac{1}{1-\|x\|^2}}, & \|x\|<1\\ 0, & \|x\|\ge 1\end{cases}$$

where θ is defined by $\int j\,dx=1$. For $\zeta,\eta>0$ set

$$j_\zeta(x)=\frac{1}{\zeta^n}j\left(\frac{x}{\zeta}\right)$$

and

$$\chi_{\eta,\zeta}(x)=c_{\eta,\zeta}j_\zeta(x)\int j_\eta(x-y)\chi(y)dy$$

with $\int\chi_{\eta,\zeta}(x)dx=1$. Then j_ζ and $\chi_{\eta,\zeta}$ are mollifiers. Moreover, for any fixed $\zeta>0$ we have

$$\lim_{\eta\to 0}c_{\eta,\zeta}=\frac{1}{\int j_\zeta(x)\chi(x)dx}>0.$$

Now,

$$D^2V_{\chi_{\eta,\zeta}}(x_0)-V_0=\int_{\|y-x_0\|\le\zeta}\chi_{\eta,\zeta}(x_0-y)(D^2V(y)-V_0)dy=$$

$$=\int_{\substack{\|y-x_0\|\le\zeta\\ y-x_0\in\mathcal{C}}}\chi_{\eta,\zeta}(x_0-y)(D^2V(y)-V_0)dy+\int_{\substack{\|y-x_0\|\le\zeta\\ y-x_0\notin\mathcal{C}}}\chi_{\eta,\zeta}(x_0-y)(D^2V(y)-V_0)dy.$$

Hence,

$$\|D^2V_{\chi_{\eta,\zeta}}(x_0)-V_0\|\le$$

$$\le\operatorname{ess}\sup_{\substack{\|y-x_0\|\le\zeta\\ y-x_0\in\mathcal{C}}}\|D^2V(y)-V_0\|+c_{\eta,\zeta}(\|D^2V(\cdot)\|_\infty+\|V_0\|)\int_{\substack{\|y\|\le\zeta\\ y\in\mathcal{C}_\eta\setminus\mathcal{C}}}j_\zeta(y)dy\le$$

$$\leq \text{ess} \sup_{\substack{\|y-x_0\|\leq\zeta \\ y-x_0\in\mathcal{C}}} \|D^2V(y)-V_0\| + c_{\eta,\zeta}(\|D^2V(\cdot)\|_\infty + \|V_0\|)\int_{\substack{\|y\|\leq 1 \\ y\in\mathcal{C}_{\eta/\zeta}\setminus\mathcal{C}}} j(y)dy, \tag{18}$$

where $\mathcal{C}_\eta$ is the η-neighbourhood of $\mathcal{C}$. Now, for ζ small enough the first term in (18) can be made arbitrarily small; for any ζ we can choose η such that the second term is arbitrarily small. Thus, we have found a sequence $\eta, \zeta \to 0$ for which $D^2V_{\chi_{\eta,\zeta}}(x_0) \to V_0$. Q.E.D.

As a consequence, in Example 2.3 at the points of discontinuity S_k may be any value between the right and the left limit of f''. The mollifying technique which we have used in our proof above can be used in the practical implementation by modifying the process itself ([9]).

3. The case of a general equation $f(x) = 0$

It would be interesting to prove a similar convergence theorem (i) in the case of an indefinite D^2V (the so-called saddle-point or mountain-pass problem) and (ii) in the case in which f does not stem from a potential V. In the first case a related condition is

$$K_2 = \sup_{\xi\neq 0,\, x,y\in\Omega} \frac{\xi^T|D^2V(x)|\xi}{\xi^T|D^2V(y)|\xi} < 2\,, \tag{19}$$

where $|A|$ means $\sqrt{A^2}$ for a selfadjoint A. In the second case the condition would be

$$K_1 = \sup_{x,y\in\Omega,\, \xi\neq 0} \frac{\|Df(x)\xi\|}{\|Df(y)\xi\|} < 2. \tag{20}$$

Note the following facts:

(1) If valid on the whole of X, this condition immediately implies the uniform boundedness of $\|Df(x)^{-1}\|$. This, however, is known to imply that f is a diffeomorphism of X onto itself (see e.g. [6] Ch. 5 and the references cited there) and the equation $f(x) = 0$ has a unique solution.

(2) (20) is stronger than (2), i.e. $K_1 \geq K$, whenever f is the gradient of a uniformly convex potential V. In the one dimensional case we have $K_1 = K$.

(3) The bound $K_1 = K = 2$ is sharp, as was shown in the one-dimensional example above.

(4) If K_1 is equal to one the function f is in fact affine, and the process (1) converges in one step.

We will now prove the last property in the finite dimensional case (the proof in the general case should be similar).

Theorem 3.1. *Let Ω be an open connected subset of R^n and let*

$$f : \Omega \to R^n$$

be continuously differentiable such that $Df(x_0)$ is non-singular for some $x_0 \in \Omega$ and that (20) holds with $K_1 = 1$. Then

$$f(x) = Ax + b,$$

where A is a fixed non-singular matrix and b is a fixed vector.

Proof. With $K_1 = 1$ we can rewrite (20) as

$$\|Df(x)\xi\| = \|Df(y)\xi\| \tag{21}$$

(the equality sign comes from the fact that x, y can be exchanged). From (21) it immediately follows that for any x, y

$$Df(y) = U(x,y)Df(x),$$

where the matrix $U(x,y)$ is orthogonal. In particular, all $Df(x)$ are non-singular. Take any $x_1 \in \Omega$. Then there is an open ball Γ at x_1 such that

$$f : \Gamma \to f(\Gamma)$$

is a diffeomorphism and a smaller Γ_1 such that

$$conv(f(\Gamma_1)) \subseteq f(\Gamma).$$

We now show that this diffeomorphism "preserves the arcs" in the following sense: if $x(t), 0 \le t \le 1$ is any continuously differentiable path in Γ_1 then the paths $Df(x_1)x(t)$ and $f(x(t))$ have the same lengthes. Indeed,

$$\int \|\frac{d}{dt}f(x(t))\|dt = \int \|Df(x(t))\dot{x}(t)\|dt =$$

$$= \int \|U(x_1, x(t))Df(x_1)\dot{x}(t)\|dt = \int \|Df(x_1)\dot{x}(t)\|dt.$$

Set now $x, y \in \Gamma_1$, $x(t) = x + t(y-x)$. Then the right hand side of the equation above equals $\|Df(x_1)(y-x)\|$ whereas the left hand side is not smaller than $\|f(y) - f(x)\|$ — the straight line connection is always the shortest. Conversely, set $f(x(t)) = f(x) + t(f(y) - f(x))$; then the left hand side integral is $\|f(y) - f(x)\|$ and the right hand side is not smaller than $\|Df(x_1)(y-x)\|$. Altogether,

$$\|f(y) - f(x)\| = \|Df(x_1)(y-x)\|.$$

In other words, the map

$$f(Df(x_1)^{-1}\cdot) : Df(x_1)\Gamma_1 \to f(\Gamma_1)$$

preserves the Euclidean length and it is well known to have the form

$$f(Df(x_1)^{-1}x) = Ux + b,$$

where U is a fixed orthogonal matrix and f is a fixed vector. This gives the assertion of the theorem with $A = UDf(x_1)$, at least on Γ_1. The extension to the whole of Ω is made by the usual connectedness argument. Q.E.D.

We next show that (2) is implied by (20). Indeed, (20) can be rewritten as

$$\xi^T D^2V(x)^2\xi \le K_1^2 \xi^T D^2V(y)^2\xi$$

for all x, y, ξ. This, however, is known to imply[3]

$$\xi^T D^2V(x)\xi \le K_1 \xi^T D^2V(y)\xi$$

[3]See e.g. [4], Ch. V, Th. 4.12.

for all x, y, ξ (but not vice versa). So, in this case the condition (20) is stronger than (2) i.e. $K \leq K_1$, except in the one-dimensional case in which $K_1 = K$. This suggests a convergence condition for the 'intermediate' case where V is indefinite and satisfies (19). Here, again, $K_2 = 1$, if and only if DV is affine. Indeed, in this case

$$|D^2V(y)| = |D^2V(x)|, \qquad \text{for all } x, y.$$

By using the polar decomposition this is immediately seen to be equivalent to

$$D^2V(y) = U(x, y)D^2V(x)$$

as in the proof of the theorem above. This indicates that the convergence could be proved under (20), (19), respectively, probably by using some more sophisticated technique. One obvious difficulty is, how to tell, whether one is approaching a saddle point. This will be the subject of our future research.

Acknowledgement. The author is indebted to A. Wiegner, Hagen for suggesting Example 2.3 above. He is very particularly indebted to an anonymous referee for his very detailed comments, which have greatly helped to give the text its present form.

References

[1] Argyros, I. K., On Newton's method under mild differentiability conditions, Appl. Math. Comp. **102** (1999) 177-183.

[2] Guttiérez, J.M., Hernández, M.A., Newton's method under weak Kantorovich conditions, IMA J. Numer. Anal. **20**, (2000) 521-532.

[3] Kantorovič, L.V., Akilov, G.P., Functional Analysis (Russian) Nauka, Moskva 1977.

[4] Kato, T., Perturbation Theory for Linear Operators. 2nd ed. Springer-Verlag, Berlin, 1976.

[5] Nesterov, Yu., Nemirovsky, A., Acceleration and parallelization of the path-following interior point method for a linearly constrained convex quadratic problem, SIAM J. Optim. **1** (1991), no. 4, 548-564.

[6] Ortega, J. M., Rheinboldt, W.C., Iterative Solution of nonlinear equations in several variables, Acad. Press NY 1970.

[7] Qi, L., Sun, J., A Nonsmooth Version of Newton's Method, Math. Programming, **58** (1993) 353-367.

[8] Sun, D., Han, J., Newton and Quasi-Newton Methods for a class of Nonsmooth Equations and Related Problems, SIAM J. Optim. **7** (1997) 463-480.

[9] Xu, H., Adaptive Smoothing Method, Deterministically Computable Generalized Jacobians and the Newton Method, J. Opt. Th. Appl. **109** (2001) 215-224.

Lehrgebiet Mathematische Physik, Fernuniversität Hagen, Postf. 940,, 58084 Hagen Germany

E-mail address: kresimir.veselic@fernuni-hagen.de