REVUE D'ANALYSE NUMÉRIQUE ET DE THÉORIE DE L'APPROXIMATION
Rev. Anal. Numér. Théor. Approx., vol. 37 (2008) no. 2, pp. 119–125
ictp.acad.ro/jnaat

NEWTON'S METHOD IN RIEMANNIAN MANIFOLDS

IOANNIS K. ARGYROS*

Abstract. Using more precise majorizing sequences than before [1], [8], and under the same computational cost, we provide a finer semilocal convergence analysis of Newton's method in Riemannian manifolds with the following advantages: larger convergence domain, finer error bounds on the distances involved, and a more precise information on the location of the singularity of the vector field.

MSC 2000. 65H10, 65B05, 65G99, 47H17, 49M15.
Keywords. Newton's method, Riemannian manifold, local/semilocal convergence, singularity of a vector field, Newton-Kantorovich method.

1. INTRODUCTION

We refer the reader to [1], [5], [7] for some of the concepts introduced but not detailed here.

Let X be a C^1 vector field defined on a connected, complete and finite-dimensional Riemannian manifold (M, g). In this study we are concerned with the following problem:

$$\text{find } p^* \in M \text{ such that } X(p^*) = o \in T_{p^*}M. \tag{1.1}$$

A point p^* satisfying (1.1) is called a singularity of X.

The most popular method for generating a sequence $\{p_n\}$ $(n \geq 0)$ approximating p^* is undoubtedly Newton's method, described here as follows:

Assume there exists an initial guess $p_0 \in M$ such that the covariant $X'(p_0)$ of X at p_0 given by

$$X'(p)v := \nabla_v X(p) = (\nabla_y X)(p), \quad v \in T_pM, \tag{1.2}$$

is invertible, at $p = p_0$, for each pair of continuously differentiable vector fields X, Y where the vector field $\nabla_y X$ stands for the covariant derivative of X with respect to Y.

Define the Riemannian-Newton method by

$$p_{n+1} = \exp_{p_n}[-X'(p_n)^{-1}X(p_n)], \tag{1.3}$$

where $\exp_p \colon T_pM \to M$ is the exponential map at p.

*Cameron university, Department of Mathematics Sciences, Lawton, OK 73505, USA, e-mail: iargyros@cameron.edu.

A survey of local as well as semilocal convergence results for method (1.3) can be found in [1]–[10] and the references there.

Here we are motivated by optimization considerations and the recent excellent study [1] of the Riemannian analogue of the property used by Zabreijko and Nguen [10].

In particular we show:

Under weaker hypotheses and the same computational cost finer error bounds on the distances $d(p_{n+1}, p_n)$, $d(p_n, p^*)$ $(n \geq 0)$ and a more precise information on the location of the solution p^* are obtained.

All the above advantages are achieved because we use more precise error estimates on the distances involved than in [1] along the lines of our relevant works for Newton's method for solving nonlinear equations on Banach spaces [2]–[6].

In Section 2 we cover the local whereas in Section 3 we study the semilocal convergence of method (1.3).

2. LOCAL CONVERGENCE ANALYSIS OF METHOD (1.3)

We need the definition [1], [5]:

Definition 2.1. *Let $G_2(p_0, r)$ denote the class of all the piecewise geodesic curves $c\colon [0,T] \to M$ which satisfy:*

(a) *$c(0) = p_0$ and the length of c is no greater than r;*
(b) *there exists $\tau \in (0,T)$ such that $c\big|_{[0,\tau]}$ is a minimizing geodesic and $c\big|_{[\tau,T]}$ is a geodesic.*

We can now introduce a Lipschitz as well as a center-Lipschitz-type continuity of X':

Let $R > 0$. We suppose there exist continuous and nondecreasing functions $\ell_0, \ell\colon [0,R] \to [0,+\infty)$ such that: for every $r \in [0,R]$ and $c \in G_2(p_0, r)$,

$$\|X'(p_0)^{-1}[P_{c,b,0}X'(c(b)) - P_{c,0,0}X'(c(0))]\|_{p_0} \leq \ell_0(r)\int_0^b |\dot{c}|,\ \ 0 \leq b, \tag{2.1}$$

$$\|X'(p_0)^{-1}[P_{c,b,0}X'(c(b)) - P_{c,a,0}X'(c(a))]\|_{p_0} \leq \ell(r)\int_a^b |\dot{c}|, \quad 0 \leq a \leq b, \tag{2.2}$$

where

$$P_{c,t,0}T(c(t)) = T(c(0)) + \int_0^t [P_{c,s,0}T'(c(s))\dot{c}(s)]\mathrm{d}s. \tag{2.3}$$

Without loss of generality we assume $\ell_0(r) > 0$, $\ell(r) > 0$ on $(0,R]$.

Remark 2.2. In general

$$\ell_0(r) \leq \ell(r), \quad r \in [0,R] \tag{2.4}$$

holds and $\frac{\ell(r)}{\ell_0(r)}$ can be arbitrarily large [3]–[6]. It is convenient to define parameter

$$\eta = |X'(p_0)^{-1}X(p_0)|_{p_0}, \tag{2.5}$$

functions $v, w\colon [0,R] \to [-\infty, +\infty)$ by

$$w(r) = \eta - r + \int_0^r (r-s)\ell(s)\mathrm{d}s, \tag{2.6}$$

$$v(r) = -1 + \ell_0(r) \tag{2.7}$$

and iterations $\{t_n\}$, $\{r_n\}$ $(n \geq 0)$ by

$$t_0 = 0, \quad t_1 = \eta, \quad t_{n+1} = t_n - \frac{w(t_n)}{v(t_n)}, \tag{2.8}$$

$$r_0 = 0, \quad r_{n+1} = r_n - \frac{w(r_n)}{w'(r_n)} \tag{2.9}$$

for all $n \geq 0$. □

Let us consider the assumption: the function w given by (2.6) has a unique zero r^* in $[0,R]$ with

$$w(R) \leq 0. \tag{2.10}$$

We showed in [2] (see also [6], [7]):

Proposition 2.3. *Under hypothesis* (2.10) *iterations* $\{t_n\}$, $\{s_n\}$ $(n \geq 0)$ *are well defined, monotonically increasing and convergent to* t^*, r^*, *respectively, with*

$$t^* \leq r^*. \tag{2.11}$$

Moreover, the following estimates hold for all $n \geq 0$

$$t_n \leq r_n, \tag{2.12}$$

$$t_{n+1} - t_n \leq r_{n+1} - r_n, \tag{2.13}$$

and

$$t^* - t_n \leq r^* - r_n. \tag{2.14}$$

Furthermore if (2.4) *holds as a strict inequality so do* (2.12) *and* (2.13) *for* $n \geq 1$. *Since we shall show both* $\{t_n\}$, $\{r_n\}$ *are majorizing sequences for* $\{p_n\}$, *it follows by Proposition* 2.3 *that the claims made in the introduction for the local convergence of method* (1.3) *hold true.*

We can now state the main local convergence result for method (1.3) which improves the corresponding Theorem 3.1 in [1, p. 8]:

Theorem 2.4. *Under hypotheses* (2.1), (2.2) *and* (2.10) *the following hold true:*

(a) *the vector field* X *admits a unique singularity* p^* *in* $\overline{U}(p_0, R) = \{p \in X \mid \|p - p_0\| \leq R\}$ *which belongs to* $\overline{U}(p_0, t^*)$. *If* $\ell_0(t^*) < 0$ *then* $X'(p^*) \in GL(T_{p^*}M)$.

(b) *Sequence* $\{p_n\}$ $(n \geq 0)$ *generated by method* (3) *is well defined,* $p_n \in \overline{U}(p_0, t_n)$ *for all* $n \geq 0$, *and* $\lim_{n\to\infty} p_n = p^*$.

(c) *The following estimates hold true for all* $n \geq 0$:

$$d(p_{n+1}, p_n) \leq |X'(p_n)^{-1}X(p_n)|_{p_n} \leq t_{n+1} - t_n \leq r_{n+1} - r_n, \tag{2.15}$$

and

$$d(p_n, p^*) \leq t^* - t_n \leq r^* - r_n. \tag{2.16}$$

Proof. Uses the more accurate (i.e. the one really needed) condition (2.1) instead of condition (2.2) used in [1] for the computation of the inverses $X'(p_n)^{-1}$ $(n \geq 0)$. The rest of the proof is identical to [1] and is omitted. □

Remark 2.5.

(a) If $\ell_0(r) = \ell(r)$ our Theorem 2.4 reduces to Theorem 3.1 in [1]. Otherwise (see (2.4)) it is an improvement over it as already shown in Proposition 2.3. Note also that the rest of the bounds obtained in Theorem 3.1 in [1] (see parts (iv)-(v) of Theorem 3.1) hold true with $\{t_n\}$ replacing sequence $\{r_n\}$ there but we decided not to include those bounds here to avoid repetitions.

(b) Theorem 2.4 remains valid for a C^1 vector field $X\colon D \subset M \to TM$ which is defined only on an open subset D of M provided $\overline{U}(p_0, R) \subseteq D$ [1], [5], [7].

(c) It follows from the proof of the theorem that the sharper (than $\{t_n\}$) scalar $\{\bar{t}_n\}$ $(n \geq 0)$ given by

$$\bar{t}_0 = 0,\ \bar{t}_1 = \eta,\ \bar{t}_{n+2} = \bar{t}_{n+1} + \frac{1}{1-\ell_0(t_{n+1})} \int_0^1 \ell(\bar{t}_n + t(\bar{t}_{n+1} - \bar{t}_n))(\bar{t}_{n+1} - \bar{t}_n)\mathrm{d}t\ (n \geq 0), \tag{2.17}$$

is also a majorizing sequence for $\{p_n\}$ $(n \geq 0)$.

Sufficient convergence conditions for sequence (1.27) which are weaker than (2.10) have already been given in [3]–[7]. Note that

$$\bar{t}_n \leq t_n, \tag{2.18}$$

$$\bar{t}_{n+1} - \bar{t}_n \leq t_{n+1} - t_n, \tag{2.19}$$

$$\bar{t}^* - \bar{t}_n \leq t^* - t_n, \tag{2.20}$$

and

$$\bar{t}^* \leq t^*. \tag{2.21}$$

See also Remark 3.3. □

3. SEMILOCAL CONVERGENCE ANALYSIS OF METHOD (1.3)

An extension of the Newton-Kantorovich theorem to finite-dimensional and complete Riemannian manifolds has been given by Ferreira and Svaiter in [7] and has been improved by us in [4]. Moreover the results in [7] were shown to be special cases of Theorem 3.1 in [1]. Here we weaken these results using L-Lipschitz and L_0-center Lipschitz conditions for tensors.

Definition 3.1. *A $(1,k)$-tensor T on M is said to be: L_0-center Lipschitz continuous on a subset S of M, if for all geodesic curve $\gamma\colon [0,1] \to M$ with endpoints in S and $p_0 \in M$*

$$\|P_{\gamma,1,0}T(\gamma(1)) - T(p_0)\|_{p_0} \leq L_0 \int_0^1 |\dot\gamma(t)| \mathrm{d}t, \tag{3.1}$$

and L-Lipschitz continuous on S, if

$$\|P_{\gamma,1,0}T(\gamma(1)) - T(\gamma(0))\|_{\gamma(0)} \leq L \int_0^1 |\dot\gamma(t)| \mathrm{d}t. \tag{3.2}$$

We can show the following improvement of Theorem 5.1 in [1] for the semilocal convergence of method (1.3) (see [3, page 387, Case 3, for $\delta = \delta_0$]):

Theorem 3.2. *Under hypotheses* (3.1) *and* (3.2) *on $S = \overline{U}(x_0, R)$ further suppose there exists $p_0 \in M$ such that $X'(p_0) \in GL(T_{p_0}M)$. Set:*

$$a = \|X'(p_0)^{-1}\|_{p_0}$$

and

$$\overline{L} = \tfrac{1}{8}\ (L + 4\ L_0 + \sqrt{L^2 + 8\ L_0\ L}).$$

Assume:

$$h_0 = a\eta\overline{L} \leq \tfrac{1}{2}\,, \tag{3.3}$$

$$\overline{U}(p_0, s^*) \subseteq \overline{U}(p_0, R), \tag{3.4}$$

where

$$s^* = \lim_{n\to\infty} s_n \leq b\ \eta, \quad b = \tfrac{2}{2-b_0}, \quad b_0 = \tfrac{1}{2}\left[-\tfrac{L}{L_0} + \sqrt{\left(\tfrac{L}{L_0}\right)^2 + 8\ \tfrac{L}{L_0}}\ \right], \tag{3.5}$$

$$s_0 = 0, \quad s_1 = \eta, \quad s_{n+2} = s_{n+1} + \tfrac{a\ L(s_{n+1}-s_n)^2}{2\ (1-L_0 s_{n+1})} \quad (n \geq 0). \tag{3.6}$$

Then

(a) *scalar sequence $\{s_n\}$ generated by* (3.6) *is monotonically increasing, bounded above by $b\ \eta$ and converges to $s^* \in [\eta, b\ \eta]$.*

(b) *Sequence $\{p_n\}$ $(n \geq 0)$ generated by method* (1.3) *is well defined, remains in $\overline{U}(p_0, s^*)$ for all $n \geq 0$ and converges to a unique singularity of X in $\overline{U}(p_0, s^*)$.*

(c) *The following error bounds hold true for all $n \geq 0$:*

$$d(p_{n+1}, p_n) \leq s_{n+1} - s_n, \tag{3.7}$$

and

$$d(p_n, p^*) \leq s^* - p_n. \tag{3.8}$$

(d) *If there exists* $R_1 \in [s^*, R]$ *such that*

$$aL_0(s^* + R_1) \leq 2, \tag{3.9}$$

then p^* *is unique in* $U(p_0, R_1)$.

Proof. As the proof in Theorem 5.1 in [1, p. 18] but we use condition (3.1) for the computation of the upper bounds of $X'(p_n)^{-1}$ $(n \geq 0)$ which is really needed instead of (3.2) used in [1]. □

Remark 3.3. (a) As in Remark 2.2

$$L_0 \leq L \tag{3.10}$$

holds in general and $\frac{L}{L_0}$ can be arbitrarily large [3]–[7]. If $L_0 = L$ holds then our theorem reduces to Theorem 5.1 in [1]. Otherwise it is an improvement. Indeed, the condition corresponding to (3.3) in [1] is given by

$$h = a\eta L \leq \tfrac{1}{2}\,. \tag{3.11}$$

Note that

$$h \leq \tfrac{1}{2} \Rightarrow h_0 \leq \tfrac{1}{2} \tag{3.12}$$

but not vice versa.

Moreover the corresponding majorizing sequence is given by

$$z_0 = 0, \quad z_{n+1} = z_n - \frac{w_1(z_n)}{w_1'(z_n)}\,, \tag{3.13}$$

where

$$w_1(r) = \eta - r + \frac{aLr^2}{2}\,, \tag{3.14}$$

and again

$$s_n \leq z_n, \tag{3.15}$$

$$s_{n+1} - s_n \leq z_{n+1} - z_n, \tag{3.16}$$

$$s^* - s_n \leq z^* - z_n, \tag{3.17}$$

and

$$s^* \leq z^* = \lim_{n\to\infty} z_n = \frac{1-\sqrt{1-2h}}{aL} \tag{3.18}$$

with (3.15), (3.16) holding as strict inequalities for $n \geq 1$ if $L_0 < L$. Finally note that in [3] we provided sufficient convergence conditions for iteration (3.6) that are even weaker than (3.3). □

All the above justify the advantages of our approach already stated in the introduction of the paper. These ideas can be used to improve the rest of the results stated in [1]. However we leave the details for the motivated reader.

REFERENCES

[1] ALVAREZ, F., BOLTE, J. and MUNIER, J., *A unifying local convergence result for Newton's method in Riemannian manifolds*, Institut National de Recherche en informatique et en automatique, Theme Num-Numeriques, Project, Sydoco, Rapport de recherche No. 5381, November 2004, France.

[2] ARGYROS, I. K., *An improved convergence analysis and applications for Newton-like methods in Banach space*, Numer. Funct. Anal. Optim., **24**, nos. 7–8, pp. 653–672, 2003.

[3] ARGYROS, I. K., *A unifying local-semilocal convergence and applications for two-point Newton-like methods in Banach space*, J. Math. Anal. Applic., **298**, pp. 374–397, 2004.

[4] ARGYROS, I. K., *On the Newton-Kantorovich method in Riemannian manifolds*, Advances in Nonlinear Variational Inequalities, **8**, no. 2, pp. 81–85, 2005.

[5] ARGYROS, I. K., *Computational theory of iterative methods*, Series: Studies in Computational Mathematics, **15**, Editors, C.K. Chui and L. Wuytack, Elsevier Publ. Co., 2007, New-York, USA.

[6] ARGYROS, I. K., *On a class of Newton–like methods for solving nonlinear equations*, J. Comput. Appl. Math., to appear.

[7] DO CARANO, M., *Riemannian Geometry*, Birkhäuser, Boston, 1992.

[8] FERREIRA, O. P. and SVAITER, B. F., *Kantorovich's theorem on Newton's method in Riemannian manifolds*, J. Complexity, **18**, pp. 304–353, 2002.

[9] KANTOROVICH, L. V. and AKILOV, G. P., *Functional Analysis in Normed Spaces*, Pergamon Press, Oxford, 1982.

[10] ZABREJKO, P. P. and NGUEN, D. F., *The majorant method in the theory of Newton-Kantorovich approximations and the Ptak error estimates*, Numer. Funct. Anal. Optim., **9**, pp. 671–674, 1987.

Received by the editors: May 12, 2008.