Applied Mathematics and Computation 261 (2015) 28–38

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Local convergence analysis of Inexact Newton method with relative residual error tolerance under majorant condition in Riemannian manifolds ☆,☆☆

Tiberio Bittencourt, Orizon Pereira Ferreira*

IME/UFG, CP-131, CEP 74001-970, Goiânia, GO, Brazil

ARTICLE INFO

Keywords:
Inexact Newton's method
Majorant principle
Local convergence analysis
Riemannian manifold

ABSTRACT

A local convergence analysis of Inexact Newton's method with relative residual error tolerance for finding a singularity of a differentiable vector field defined on a complete Riemannian manifold, based on majorant principle, is presented in this paper. We prove that under local assumptions, the Inexact Newton method with a fixed relative residual error tolerance converges Q linearly to a singularity of the vector field under consideration. Using this result we show that the Inexact Newton method to find a zero of an analytic vector field can be implemented with a fixed relative residual error tolerance. In the absence of errors, our analysis retrieves the classical local theorem on the Newton method in Riemannian context.

1. Introduction

Newton's method and its variations, including the Inexact Newton methods, are the most efficient methods known for solving nonlinear equations in Banach spaces. Besides its practical applications, Newton's method is also a powerful theoretical tool with a wide range of applications in pure and applied mathematics, see [6,11,16,23,24,26,33,34]. In particular, Newton's method has been instrumental in the modern complexity analysis of the solution of polynomial or analytical equations [6,25], linear and quadratic programming problems and linear semi-definite programming problems [15,16,23,24]. In all these applications, homotopy methods are combined with Newton's method, which helps the algorithm to keep track of the solution of a parameterized perturbed version of the original problem.

In classic Newton's method, a linear equation system is solved in each iteration which can be expensive and unnecessary when the problem size is large. Inexact Newton's method comes up to overcome such drawback and can effectively cut down the computational cost by solving the linear equations approximately, see [9,12,21]. It would be most desirable to have an *a priori* prescribed residual error tolerance in the iterative solutions of linear system for computing the Inexact Newton steps, in order to avoid under-solving or over-solving the linear system in question. The advantage of working with an error tolerance on the residual rests in the fact that the exact Newton step need not to be known for evaluating this error, which makes this criterion attractive for practical applications, see [15,16].

☆ This author was supported by CAPES.
☆☆ This author was supported by CNPq Grants Universal-FAPEG, CNPq Grants 471815/2012-8, 305158/2014-7 and PRONEX–Optimization (FAPERJ/CNPq).
* Corresponding author. Tel.: +55 62 3521 1424; fax: +55 62 3521 1180.
E-mail address: orizon@ufg.br, orizonferreira@gmail.com (O.P. Ferreira).

http://dx.doi.org/10.1016/j.amc.2015.03.080
0096-3003/

Newton's method has been extended to Riemannian manifolds with many different purposes. In particular, in the last few years, a couple of papers have dealt with the issue of convergence analysis of Newton's method for finding a singularity of a differentiable vector field defined on a complete Riemannian manifold, see [1–5,8,13,14,18–20,26,28–32]. Extensions to Riemannian manifolds of analyses of Newton's method under the γ-condition was given in [8,18–20]. Although the local convergence analysis of Inexact Newton's method in Banach space with relative errors tolerance in the residue [7,9,21] are well understood, as far as we know, the convergence analysis of the method in Riemannian manifolds context under general local assumptions, assuming *only* bounded relative residual errors, is a new contribution of this paper. It is worth to point out that, for null error tolerance, the analysis presented merge in the usual local convergence analysis on Newton's method in Riemannian manifold under a majorant condition, see [13]. In our analysis, the classical Lipschitz condition is relaxed using a majorant function which provides a clear relationship between the majorant function and the vector field under consideration. Moreover, several unrelated previous results pertaining to Newton's method are unified (see [8,18,19]), now in the Riemannian context.

The organization of the paper is as follows. In Section 2, the notations and basic results used in the paper are presented. In Section 3 the main result is stated and in Section 4 some properties of the majorant function are established and the main relationships between the majorant function and the vector field used in the paper are presented. In Section 5 the main result is proved and two applications of this result are given in Section 6. Some final remarks are made in Section 7.

2. Notation and auxiliary results

In this section we recall some notations, definitions and basic properties of Riemannian manifolds used throughout the paper, they can be found, for example in [10] and [17].

Throughout the paper, $\mathcal{M}$ is a smooth manifold and $C^1(\mathcal{M})$ is the class of all continuously differentiable functions on $\mathcal{M}$. The space of vector fields on $\mathcal{M}$ is denoted by $\mathcal{X}(\mathcal{M})$, by $T_p\mathcal{M}$ we denote the tangent space of $\mathcal{M}$ at p and by $T\mathcal{M} = \bigcup_{x\in\mathcal{M}} T_x\mathcal{M}$ the *tangent bundle* of $\mathcal{M}$. Let $\mathcal{M}$ be endowed with a Riemannian metric $\langle\cdot,\cdot\rangle$, with corresponding norm denoted by $\|\cdot\|$, so that $\mathcal{M}$ is now a *Riemannian manifold*. Let us recall that the metric can be used to define the length of a piecewise C^1 curve $\zeta : [a,b] \to \mathcal{M}$ joining p to q, i.e., such that $\zeta(a) = p$ and $\zeta(b) = q$, by $l(\zeta) = \int_a^b \|\zeta'(t)\|dt$. Minimizing this length functional over the set of all such curves we obtain a distance $d(p,q)$, which induces the original topology on M. The open and closed balls of radius $r > 0$ centered at p are defined, respectively, as

$$B_r(p) := \{q \in M : d(p,q) < r\}, \quad \bar{B}_r(p) := \{q \in M : d(p,q) \le r\}.$$

Also the metric induces a map $f \in C^1(\mathcal{M}) \mapsto \operatorname{grad} f \in \mathcal{X}(\mathcal{M})$, which associates to each f its *gradient* via the rule $\langle \operatorname{grad} f, X\rangle = df(X)$, for all $X \in \mathcal{X}(\mathcal{M})$. The chain rule generalizes to this setting in the usual way: $(f\circ\zeta)'(t) = \langle \operatorname{grad} f(\zeta(t)), \zeta'(t)\rangle$, for all curves $\zeta \in C^1$. Let ζ be a curve joining the points p and q in $\mathcal{M}$ and let ∇ be a Levi–Civita connection associated to $(\mathcal{M}, \langle,\rangle)$. For each $t \in [a,b]$, ∇ induces an isometry, relative to $\langle\cdot,\cdot\rangle$,

$$\begin{aligned} P_{\zeta,a,t} : T_{\zeta(a)}\mathcal{M} &\longrightarrow T_{\zeta(t)}\mathcal{M} \\ v &\longmapsto P_{\zeta,a,t}v = V(t), \end{aligned} \tag{1}$$

where V is the unique vector field on ζ such that $\nabla_{\zeta'(t)}V(t) = 0$ and $V(a) = v$, the so-called *parallel translation* along ζ from $\zeta(a)$ to $\zeta(t)$. Note also that

$$P_{\zeta,b_1,b_2} \circ P_{\zeta,a,b_1} = P_{\zeta,a,b_2}, \quad P_{\zeta,b,a} = P_{\zeta,a,b}{}^{-1}.$$

A vector field V along ζ is said to be *parallel* if $\nabla_{\zeta'}V = 0$. If ζ' itself is parallel, then we say that ζ is a *geodesic*. The geodesic equation $\nabla_{\zeta'}\zeta' = 0$ is a second order nonlinear ordinary differential equation, so the geodesic ζ is determined by its position p and velocity v at p. It is easy to check that $\|\zeta'\|$ is constant. We say that ζ is *normalized* if $\|\zeta'\| = 1$. A geodesic $\zeta : [a,b] \to \mathcal{M}$ is said to be *minimal* if its length is equal the distance of its end points, i.e. $l(\zeta) = d(\zeta(a), \zeta(b))$.

A Riemannian manifold is *complete* if its geodesics are defined for any values of t. The Hopf–Rinow's theorem asserts that if this is the case then any pair of points, say p and q, in $\mathcal{M}$ can be joined by a (not necessarily unique) minimal geodesic segment. Moreover, $(\mathcal{M}, d)$ is a complete metric space and bounded and closed subsets are compact.

The *exponential map* at p, $\exp_p : T_p\mathcal{M} \to \mathcal{M}$ is defined by $\exp_p v = \zeta_v(1)$, where ζ_v is the geodesic defined by its position p and velocity v at p and $\zeta_v(t) = \exp_p tv$ for any value of t. For $p \in \mathcal{M}$, let

$$r_p := \sup\left\{r > 0 : \exp_{p}\big|_{B_r(o_p)} \text{ is a diffeomorphism}\right\},$$

where o_p denotes the origin of $T_p\mathcal{M}$ and $B_r(o_p) := \{v \in T_p\mathcal{M} : \|v - o_p\| < r\}$. Note that if $0 < \delta < r_p$ then $\exp_p B_\delta(o_p) = B_\delta(p)$. The number r_p is called the *injectivity radius* of $\mathcal{M}$ at p.

Definition 1. Let $p \in \mathcal{M}$ and r_p the radius of injectivity at p. Define the quantity

$$K_p := \sup\left\{\frac{d(\exp_q u, \exp_q v)}{\|u - v\|} : q \in B_{r_p}(p),\, u, v \in T_q\mathcal{M},\, u \ne v,\, \|v\| \le r_p,\, \|u - v\| \le r_p\right\}.$$

Remark 1. The quantity K_p measures how fast the geodesics spread apart in $\mathcal{M}$. In particular, when $u = 0$ or more generally when u and v are on the same line through o_q,

$$d(\exp_q u, \exp_q v) = \|u - v\|.$$