Applied Mathematics and Computation 227 (2014) 762–787

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

Newton-type methods on Riemannian manifolds under Kantorovich-type conditions

S. Amat [a,1], I.K. Argyros [b,*], S. Busquier [a,1], R. Castro [c,2], S. Hilout [d,e], S. Plaza [c,2]

[a] Dpto.de Matemática Aplicada y Estadística, Universidad Politécnica de Cartagena, Spain
[b] Cameron University, Department of Mathematical Sciences, Lawton, OK 73505, USA
[c] Departamento de Matemáticas, Facultad de Ciencias, Universidad de Valparaiso, Chile
[d] UQÀM, Département des Sciences de la Terre et de l'Atmosphère, C.P. 8888, Succursale Centre-ville Montréal, Québec H3C 3P8, Canada
[e] Poitiers University, Laboratoire de Mathématiques et Applications, 86962 Futuroscope Chasseneuil Cedex, France

ARTICLE INFO

Keywords:
Third-order iterative methods
Riemannian manifolds
Kantorovich conditions
Semilocal convergence
Order of convergence

This paper is dedicated to the memory of Sergio Plaza.

ABSTRACT

One of the most studied problems in numerical analysis is the approximation of nonlinear equations using iterative methods. In the last years, attention has been paid in studying Newton's method on manifolds. In this paper, we generalize this study considering some Newton-type iterative methods. A characterization of the convergence under Kantorovich type conditions and optimal estimates of the error are found. Using normal coordinates the order of convergence is derived. The sufficient semilocal convergence criteria are weaker and the majorizing sequences are tighter for the special cases of simplified Newton and Newton methods than in earlier studies such as Argyros (2004, 2007, 2008) [6,8,12] and Kantorovich and Akilov (1964) [32].

1. Introduction

Let us suppose that F is an operator defined on an open convex subset Ω of a Banach space E. Let us denote by $DF(x_n)$ the first Fréchet derivatives of F at x_n.

Given an integer m and an initial point $x_0 \in E$, we move from x_n to x_{n+1} through an intermediate sequence $\{y_n^i\}_{i=0}^m,\ y_n^0 = x_n$, which is a generalization of Newton ($m = 1$) and simplified Newton ($m = \infty$) methods

$$\begin{cases} y_n^1 = y_n^0 - \mathcal{D}F(y_n^0)^{-1}F(y_n^0), \\ y_n^2 = y_n^1 - \mathcal{D}F(y_n^0)^{-1}F(y_n^1), \\ \vdots \\ y_n^m = x_{n+1} = y_n^{m-1} - \mathcal{D}F(y_n^0)^{-1}F\left(y_n^{(m-1)}\right). \end{cases}$$

This family of methods was introduced by Shamanskii [43]. Under appropriate conditions, these iterative methods converge to a root x_* of the equation $F(x) = 0$. Moreover, if x_0 is sufficiently near x_* the method has order of convergence at least $m + 1$. See [33,38,43,46]. In particular, Notice that in [38] a modification of $\mathcal{D}F(x_n)$ at each sub-step. In [39–41], Parida

* Corresponding author.
E-mail addresses: sergio.amat@upct.es (S. Amat), iargyros@cameron.edu (I.K. Argyros), sonia.busquier@upct.es (S. Busquier), rodrigo.castro@uv.cl (R. Castro), hilout.hilout@uqam.ca (S. Hilout).
[1] Research supported by MICINN-FEDER MTM2010-17508 (Spain) and by 08662/PI/08 (Murcia).
[2] Research supported by Fondecyt 10950252 (Chile).

0096-3003/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.amc.2013.11.055

and Gupta provided some recurrence relations to establish a convergence analysis for a third order Newton-type methods under Lipschitz or Hölder conditions on the second Fréchet derivative. A modification of the approach used in [39] and some applications are presented by Chun et al. [19]. Recently, Argyros and Ren [17] expanded the applicability of Halley's method using a center-Lipschitz condition on the second Fréchet derivative instead of Lipschitz's condition.

On the other hand, in the last years, attention has been paid in studying Newton's method on manifolds, since there are many numerical problems posed on manifolds that arise naturally in many contexts. Some examples include eigenvalue problems, minimization problems with orthogonality constraints, optimization problems with equality constraints, invariant subspace computations. See for instance [1–3,7,15,20,21,27,29,35,36,48,49]. For these problems, one has to compute solutions of equations or to find zeros of a vector field on Riemannian manifolds.

The study about convergence matter of iterative methods is usually centered on two types: semilocal and local convergence analysis. The semilocal convergence matter is, based on the information around an initial point, to give criteria ensuring the convergence of iterative methods; while the local one is, based on the information around a solution, to find estimates of the radii of convergence balls. There is a plethora of studies on the weakness and/or extension of the hypothesis made on the underlying operators; see for example [4–6,12,14,32,34,48,49].

The semilocal convergence analysis of Newton's method is based on celebrated Kantorovich theorem [10,11,23–26,30,31,42]. This theorem is a fundamental result in numerical analysis, e.g., for providing an iterative method for computing zeros of polynomials or of systems of nonlinear equations. Moreover, this theorem is a very usefull result in nonlinear functional analysis, e.g., for establishing that a nonlinear equation in an abstract space has a solution. Let us recall Kantorovich's theorem in a Banach space setting.

Theorem 1 [32]. *Let E be a Banach space, $\Omega \subseteq E$ be an open convex set, $F : \Omega \longrightarrow \Omega$ be a continuous operator, such that, $F \in C^1$ and DF is Lipschitz on Ω*

$$\|DF(x) - DF(y)\| \leqslant l\|x - y\|, \text{ for all } x, y \in \Omega, l > 0.$$

Suppose that for some $x_0 \in \Omega$, $DF(x_0)$ is invertible and that for some $a > 0$ and $b \geqslant 0$:

$$\|DF(x_0)^{-1}\| \leqslant a,$$

$$\|DF(x_0)^{-1}F(x_0)\| \leqslant b,$$

$$h = a\,b\,l \leqslant \frac{1}{2} \tag{1}$$

and

$$B(x_0, t_*) \subseteq \Omega \text{ where } t_* = \frac{1}{a\,l}\left(1 - \sqrt{1 - 2\,h}\right).$$

If

$$v_k = -DF(x_k)^{-1}F(x_k),$$
$$x_{k+1} = x_k + v_k.$$

Then $\{x_k\}_{k\in\mathbb{N}} \subseteq B(x_0, t_*)$ and $x_k \longrightarrow p_*$, which is the unique zero of F in $B[x_0, t_*]$. Furthermore, if $h < \frac{1}{2}$ and $B(x_0, r) \subseteq \Omega$ with

$$t_* < r \leqslant t_{**} = \frac{1}{a\,l}\left(1 + \sqrt{1 - 2\,h}\right),$$

then p_* is also the unique zero of F in $B(x_0, r)$. Also, the error bound is:

$$\|x_k - x_*\| \leqslant (2\,h)^{2^k}\frac{b}{h}; \quad k = 1, 2, \ldots$$

Although the concepts will be defined later on, to extend the method on Riemannian manifolds, preliminarily we will say that the derivative of F at x_n is replaced by the covariant derivative of X at p_n:

$$\begin{aligned}\nabla_{(\cdot)}X(p_n) : \quad & T_{p_n}M \longrightarrow T_{p_n}M \\ & v \longrightarrow \nabla_Y X,\end{aligned}$$

where Y is a vector field satisfying $Y(p) = v$. We adopt the notation $\mathcal{D}X(p)\,v = \nabla_y X(p)$; hence $\mathcal{D}X(p)$ is a linear mapping of T_pM into T_pM. So, in this new context

$$-F'(x_n)^{-1}F(x_n)$$

is written as

$$-\mathcal{D}X(p_n)^{-1}X(p_n)$$