Journal of Computational and Applied Mathematics 255 (2014) 106–121

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

Third-order methods on Riemannian manifolds under Kantorovich conditions

S. Amat [a,*], S. Busquier [a], R. Castro [b], S. Plaza [b,1]

[a] Departamento de Matemática Aplicada y Estadística, Universidad Politécnica de Cartagena, Spain

[b] Departamento de Matemáticas, Facultad de Ciencias, Universidad de Valparaiso, Chile

ARTICLE INFO

Article history:
Received 23 February 2012
Received in revised form 5 April 2013

This paper is dedicated to the memory of Sergio Plaza

Keywords:
Third-order iterative methods
Riemannian manifolds
Semilocal convergence

ABSTRACT

One of the most studied problems in numerical analysis is the approximation of nonlinear equations using iterative methods. In the past years, attention has been paid in studying Newton's method on manifolds. In this paper, we generalize this study by considering a general class of third-order iterative methods. A characterization of the convergence under Kantorovich type conditions and optimal error estimates is found.

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 $L_F(x_n)$ the operator

$$L_F(x_n) = DF(x_n)^{-1} D^2F(x_n) DF(x_n)^{-1} F(x_n),$$

where $DF(x_n)$ and $D^2F(x_n)$ denote respectively the first and second Fréchet derivatives of F at x_n.

Some of the most famous third order iterative methods to find a root of the equation $F(x) = 0$ are:

- *Halley* [1–7]:

$$x_{n+1} = x_n - \left(I + \frac{1}{2}L_F(x_n)\right)^{-1} DF(x_n)^{-1} F(x_n).$$

- *Super-Halley* [1–3,5,6,8,9]:

$$x_{n+1} = x_n - \left(I + \frac{1}{2}(I - L_F(x_n))^{-1} L_F(x_n)\right) DF(x_n)^{-1} F(x_n).$$

- *Chebyshev* [1–3,10,5–7]:

$$x_{n+1} = x_n - \left(I + \frac{1}{2}L_F(x_n)\right) DF(x_n)^{-1} F(x_n).$$

* Corresponding author. Tel.: +34 968 325651; fax: +34 968 32 5694.
E-mail addresses: sergio.amat@upct.es (S. Amat), sonia.busquier@upct.es (S. Busquier), rodrigo.castrom@uv.cl, rocacastro@hotmail.com (R. Castro).

1 Deceased author.

0377-0427/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cam.2013.04.023

- *Chebyshev like methods* [1,2,10,5,6]:

$$x_{n+1} = x_n - \left(I + \frac{1}{2} L_{F(x_n)} + \lambda L_F^2 (x_n)\right) DF (x_n)^{-1} F (x_n), \quad \text{for } 0 \le \lambda \le 2.$$

- *Chebyshev–Halley methods* [1–3,5–7]:

$$x_{n+1} = x_n - \frac{1}{2} \left[I - \frac{\lambda}{2} L_F (x_n)\right]^{-1} L_F (x_n) F' (x_n)^{-1} F (x_n),$$

where I is the identity operator in E.

A review to the amount of literature on high order iterative methods in Banach spaces, in the past two decades (see the above references) shows the importance of high order iterative schemes in the approximation of nonlinear equations. The above third order methods present the difficulty to evaluate the second order Fréchet derivative. For a nonlinear system of m equations and m unknowns, the first Fréchet derivative is a matrix with m^2 entries, while the second Fréchet derivative has m^3 entries. This implies a huge amount of operations, but, in some cases, it pays to be a little more elaborated.

In this paper, we start with the general family of third-order methods

$$u_n = F' (x_n)^{-1} F (x_n)$$
$$T_n = \frac{1}{2} F' (x_n)^{-1} F'' (x_n) u_n$$
$$x_{n+1} = x_n - \left(I + T_n + \sum_{k \ge 2} \beta_k T_n^k\right) u_n,$$

where $\{\beta_k\}_{k\ge2}$ is a real decreasing sequence of positive real numbers with

$$\sum_{k\ge2} \beta_k \theta_n^k \le d \left|\theta_n^2\right|; \quad 0 \le d \le 2.$$

The convergence of this family has been studied in [1,6,11] and includes as particular cases all the mentioned third order methods.

In this paper, we extend the above general family of third-order methods to the case of equations on Riemannian manifolds. We study the convergence of the family under Kantorovich type conditions and we present optimal estimates of the error.

Finally, we would like to mention that in the past 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. Examples include eigenvalue problems, minimization problems with orthogonality constraints, optimization problems with equality constraints, and invariant subspace computations (see for instance [12–27] and the references therein).

2. Basic definitions and preliminary results

In this section, we introduce some definitions and fundamental properties of Riemannian manifolds.

Definition 1. A differentiable manifold of dimension m is a set M and a family of injective mappings $x_\alpha : U_\alpha \subset \mathbb{R}^m \longrightarrow M$ of open sets U_α of $\mathbb{R}^m$ into M such that:

(i) $\cup_\alpha x_\alpha (U_\alpha) = M$;
(ii) for any pair α, β with $x_\alpha (U_\alpha) \cap x_\beta (U_\beta) = W \ne \emptyset$, the sets $x_\alpha^{-1}(W)$ and $x_\beta^{-1}(W)$ are open sets in $\mathbb{R}^m$ and the mappings $x_\beta^{-1} \circ x_\alpha$ are differentiable;
(iii) the family $\{(U_\alpha, x_\alpha)\}$ is maximal relative to the conditions (i) and (ii).

The pair (U_α, x_α) (or the mapping x_α) with $p \in x_\alpha (U_\alpha)$ is called a parametrization (or system of coordinates) of M at p; $x_\alpha (U_\alpha)$ is then called neighborhood at p and $\left(x_\alpha (U_\alpha), x_\alpha^{-1}\right)$ is called a coordinate chart. A family $\{(U_\alpha, x_\alpha)\}$ satisfying (i) and (ii) is called a differentiable structure on M.

Let M be a real manifold, $p \in M$ and denote by T_pM the *tangent space* at p to M. Let $x : U \subset \mathbb{R}^m \longrightarrow M$ be a system of coordinates around p with $x (x_1, x_2, \ldots, x_m) = p$ and its associated basis

$$\left\{ \left.\frac{\partial}{\partial x_1}\right|_p, \left.\frac{\partial}{\partial x_2}\right|_p, \ldots, \left.\frac{\partial}{\partial x_m}\right|_p \right\}$$

in T_pM. The *tangent bundle* TM is defined as

$$TM = \left\{(p, v) ; p \in M \text{ and } v \in T_pM\right\} = \bigcup_{p\in M} T_pM$$

and provides a differentiable structure of dimension $2m$ [28]. Next, we define the concept of a Riemannian metric.