Applied Mathematics and Computation 338 (2018) 421–431

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

Multipoint secant and interpolation methods with nonmonotone line search for solving systems of nonlinear equations

Oleg Burdakov[a,*], Ahmad Kamandi[b]

[a] Department of Mathematics, Linköping University, Linköping SE-58183, Sweden
[b] Department of Mathematics, University of Science and Technology of Mazandaran, PO Box 48518-78195, Behshahr, Iran

ARTICLE INFO

MSC:
65H10
65H20
65K05

Keywords:
Systems of nonlinear equations
Quasi-Newton methods
Multipoint secant methods
Interpolation methods
Global convergence
Superlinear convergence

ABSTRACT

Multipoint secant and interpolation methods are effective tools for solving systems of nonlinear equations. They use quasi-Newton updates for approximating the Jacobian matrix. Owing to their ability to more completely utilize the information about the Jacobian matrix gathered at the previous iterations, these methods are especially efficient in the case of expensive functions. They are known to be local and superlinearly convergent. We combine these methods with the nonmonotone line search proposed by Li and Fukushima (2000), and study global and superlinear convergence of this combination. Results of numerical experiments are presented. They indicate that the multipoint secant and interpolation methods tend to be more robust and efficient than Broyden's method globalized in the same way.

1. Introduction

Consider the problem of solving a system of simultaneous nonlinear equations

$$F(x) = 0, \tag{1}$$

where the mapping $F : \mathbb{R}^n \to \mathbb{R}^n$ is assumed to be continuously differentiable. Numerical methods aimed at iteratively solving this problem are discussed in [1–3]. We focus here on those which generate iterates by the formula

$$x_{k+1} = x_k + \lambda_k p_k, \quad k = 0, 1, \dots, \tag{2}$$

where the vector $p_k \in \mathbb{R}^n$ is a search direction, and the scalar λ_k is a step length. Denote $F_k = F(x_k)$ and $F'_k = F'(x_k)$. In the Newton-type methods, the search direction has the form

$$p_k = -B_k^{-1} F_k.$$

Here the matrix $B_k \in \mathbb{R}^{n \times n}$ is either the Jacobian F'_k (Newton's method) or some approximation to it (quasi-Newton methods). For quasi-Newton methods, we consider Broyden's method [4], multipoint secant methods [5–7] and interpolation methods [8,9].

* Corresponding author.
E-mail addresses: oleg.burdakov@liu.se (O. Burdakov), ahmadkamandi@mazust.ac.ir (A. Kamandi).

https://doi.org/10.1016/j.amc.2018.05.041

Newton's method is known to attain a local quadratic rate of convergence, when $\lambda_k = 1$ for all k. The quasi-Newton methods do not require computation of any derivatives, and their local rate of convergence is superlinear.

The Newton search direction $p_k^N = -(F_k')^{-1}F_k$ is a descent direction for $\|F(x)\|$ in any norm. Moreover, as it was shown in [10,11], there exists a directional derivative of $\|F(x)\|$ calculated by the formula:

$$\|F(x_k + \lambda p_k^N)\|'_{\lambda=+0} = -\|F_k\|,$$

which is valid for any norm, even if $\|F(x)\|$ is not differentiable in x_k. This property of the Newton search direction provides the basis for constructing various backtracking line search strategies [2,3,10] aimed at making Newton's method globally convergent. An important feature of such strategies is that $\lambda_k = 1$ is accepted for all sufficiently large k, which allows them to retain the high local convergence rate of the Newton method.

In contrast to Newton's method, the search directions generated by the quasi-Newton methods are not guaranteed to be descent directions for $\|F(x)\|$. This complicates the globalization of the latter methods.

The earliest line search strategy designed for globalizing Broyden's method is due to Griewank [12]. Its drawback, as indicated in [13], is related to the case when p_k is orthogonal, or close to orthogonal, to the $\nabla\|F(x_k)\|^2$. Here and later, $\|\cdot\|$ stands for the Euclidean vector norm and the induced matrix norm. The Frobenius matrix norm will be denoted by $\|\cdot\|_F$.

Li and Fukushima [13] developed a new backtracking line search for Broden's method and proved its global superlinear convergence. In this line search, the function $\|F_k\|$ may not monotonically decrease with k. Its important feature is that it is free of the aforementioned drawback of the line search proposed in [12].

The purpose of this paper is to extend the Li–Fukushima line search to the case of the multipoint secant and interpolation methods, theoretically study their global convergence and also explore their practical behavior in numerical experiments. We are also aimed at demonstrating a higher efficiency of these methods as compared with Broyden's method in the case of expensive function evaluations.

The paper is organized as follows. In the next section, we describe the multipoint secant and interpolation methods and discuss their properties. A combination of these methods with the Li–Fukushima line search is presented in Section 3. In Section 4, we show a global and superlinear convergence of this combination. Results of numerical experiments are reported and discussed in Section 5. Finally, some conclusions are included in the last section of the paper.

2. Quasi-Newton updates

The class of quasi-Newton updates that we consider here has the form

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k)c_k^T}{s_k^T c_k}, \tag{3}$$

where $s_k = x_{k+1} - x_k$, $y_k = F_{k+1} - F_k$, and $c_k \in \mathbb{R}^n$ is a parameter.

One of the most popular quasi-Newton method of solving (1) is due to Broyden [4]. It corresponds to the choice $c_k = s_k$ and satisfies the, so-called, *secant equation*:

$$B_{k+1}s_k = y_k. \tag{4}$$

It indicates that B_{k+1} provides an approximation of the Jacobian matrix along the direction s_k. Though such an approximation is provided by B_k along s_{k-1}, it is not guaranteed that B_{k+1} retains this property because, in general, $B_{k+1}s_{k-1} \neq y_{k-1}$.

Gay and Schnabel [5] proposed a quasi-Newton updating formula of the form (3) with the aim to preserve the secant equations satisfied at some previous iterations. The resulting Jacobian approximation satisfies the following *multipoint secant equations*:

$$B_{k+1}s_i = y_i, \quad \forall i \in T_{k+1}, \tag{5}$$

where $T_{k+1} = \{i : m_k \le i \le k\}$ and $0 \le m_k \le k$. To guarantee this, the parameter in (3) is calculated by the formula

$$c_k = s_k - P_k s_k, \tag{6}$$

where $P_k \in \mathbb{R}^{n\times n}$ is an orthogonal projector on the subspace generated by the vectors $s_{m_k}, s_{m_k+1}, \ldots, s_{k-1}$, and P_k vanishes when $m_k = k$. To ensure a local superlinear convergence and stable approximation of the Jacobian, it is required in [5] that there exists $\bar{\sigma} \in (0,1)$ such that

$$\|c_k\| \ge \bar{\sigma}\|s_k\|, \quad \forall k \ge 0. \tag{7}$$

To meet this requirement, m_k is chosen as follows. If the trial choice of $m_k = m_{k-1}$ fails to satisfy (7), the vectors $s_{m_{k-1}}, \ldots, s_k$ are considered as close to linear dependent, and then a *restart* is performed by setting $m_k = k$, or equivalently, $T_{k+1} = \{k\}$. Otherwise, the trial choice is accepted, in which case the set T_{k+1} is obtained by adding $\{k\}$ to T_k.