Journal of the Operations Research
Society of Japan
2003, Vol. 46, No. 1, 1-21

SUPERLINEAR CONVERGENCE OF THE SHENG-ZOU-BROYDEN METHOD FOR NONLINEAR LEAST SQUARES PROBLEMS

Hideho Ogasawara Hiroshi Yabe
Tokyo University of Science

(Received Received July 27, 2001; Revised April 15, 2002)

Abstract We are concerned with nonlinear least squares problems. It is known that structured quasi-Newton methods perform well for solving these problems. In this strategy, two kinds of factorized structured quasi-Newton methods have been independently proposed by Yabe and Takahashi (1988), and Sheng and Zou (1988). Sheng and Zou introduced a BFGS-like update by considering how the normal equation based on an affine model may consist with the Newton equation, and dealt with a hybrid method that combines the Gauss-Newton method and their BFGS-like method. In this paper, we deal with the Sheng-Zou-Broyden family proposed by Yabe (1993), which is an extension of the update of Sheng and Zou to the Broyden-like family. Local and q-superlinear convergence of the method with this family is established for nonzero residual problems.

Keywords: Nonlinear programming, nonlinear least squares problems, Sheng-Zou-Broyden method, factorized structured Broyden updates, local and q-superlinear convergence

1. Introduction

Consider the nonlinear least squares problem

$$\text{minimize } f(x) \equiv \frac{1}{2}\sum_{i=1}^{m} r_i(x)^2 = \frac{1}{2}\|r(x)\|^2, \quad x \in \mathbf{R}^n, \tag{1.1}$$

where the residual function $r(x) \equiv (r_1(x), \dots, r_m(x))^\top$ $(m \geq n)$ is smooth and nonlinear, and superscript $\top$ denotes transpose. The gradient vector of f at x is

$$g(x) \equiv \nabla f(x) = J(x)^\top r(x)$$

and the Hessian matrix of f at x is given by

$$G(x) \equiv \nabla^2 f(x) = J(x)^\top J(x) + S(x)$$

with

$$S(x) \equiv \sum_{i=1}^{m} r_i(x)\nabla^2 r_i(x),$$

where $J(x)$ denotes the $m \times n$ Jacobian matrix of r at x whose (i,j)th element is $\partial r_i(x)/\partial x_j$. Among many numerical methods for solving (1.1), *structured* quasi-Newton methods are regarded as be the most efficient ones. These methods exploit the special structure of the Hessian matrix of f.

In this paper, we only consider the framework of the line search strategy, which has the advantage of easier implementation than the trust region strategy. Assume that there exists a local minimizer x_* of f. Given an initial estimate x_0 of x_* and an initial symmetric matrix A_0, a sequence of iterates $\{x_k\}$ is generated by the formula

$$x_{k+1} = x_k + \alpha_k d_k, \quad k = 0, 1, 2, \ldots,$$

where α_k is a suitable stepsize and d_k is a search direction given by the solution of the linear system

$$(J_k^\top J_k + A_k)d_k = -J_k^\top r_k. \tag{1.2}$$

Here $J_k = J(x_k)$, $r_k = r(x_k)$, and A_k is intended to be an approximation to $S_k = S(x_k)$, the second portion of the Hessian of f at x_k. The matrix A_k is updated so that A_{k+1} satisfies the secant condition for the whole matrix,

$$(J_{k+1}^\top J_{k+1} + A_{k+1})s_k = y_k^\sharp, \tag{1.3}$$

where

$$s_k = x_{k+1} - x_k, \quad y_k^\sharp = J_{k+1}^\top J_{k+1} s_k + y_k^\flat,$$

and $y_k^\flat$ is an approximation to $S_{k+1}s_k$. For this choice of $y_k^\sharp$, thanks to cancellation from both-hand sides, we see that the secant condition (1.3) for A_{k+1} reduces to

$$A_{k+1}s_k = y_k^\flat.$$

A historical survey on structured quasi-Newton methods can be found in Dennis, Martínez and Tapia [5].

From a computational point of view, it is desirable that the coefficient matrix of (1.2) should be positive definite in order to ensure that the solution d_k is a descent direction for f. However, since the second-order term S_k to be approximated is generally indefinite, it is not easy to construct an updating formula for A_k such that the matrix $J_k^\top J_k + A_k$ maintains positive definiteness. A possible obvious choice of A_k is a positive definite matrix. But such A_k would not be a good approximation of S_k unless S_k is positive definite.

As a remedy of this deficiency, Yabe and Takahashi [16] adopted an alternative approach which utilizes the factorized structure of $J_k^\top J_k$. Specifically, they proposed approximating the Hessian matrix $G_k = G(x_k)$ in factorized form as $(J_k + W_k)^\top (J_k + W_k)$, where the matrix W_k is a correction matrix such that $J_k^\top W_k + W_k^\top J_k + W_k^\top W_k$ is an approximation to S_k. Then a search direction d_k is given by solving the linear system

$$(J_k + W_k)^\top (J_k + W_k)d_k = -J_k^\top r_k. \tag{1.4}$$

The calculated solution d_k will be a descent direction for f unless $J_k + W_k$ is rank-deficient. The matrix W_k is updated so that W_{k+1} satisfies the secant condition

$$(J_{k+1} + W_{k+1})^\top (J_{k+1} + W_{k+1})s_k = y_k^\sharp,$$

where $y_k^\sharp = J_{k+1}^\top J_{k+1}s_k + y_k^\flat$, and $y_k^\flat$ is again an approximation to $S_{k+1}s_k$.

Yabe and Takahashi [17] proposed the factorized BFGS- and DFP-like updates, and they proved local and q-superlinear convergence of these methods. Yabe and Yamaki [19] derived a factorized structured Broyden family of updates for W_k which includes the factorized BFGS- and DFP-like updates as special cases, and extended the convergence result to this family. Xu, Ma and Kong [13] independently derived a class of factorized structured quasi-Newton updates involving the factorized BFGS- and DFP-like updates proposed by Yabe and Takahashi [16], and showed local and q-superlinear convergence of their methods. Ogasawara [9] proved that the class of updates proposed by Yabe and Yamaki [19] and the class of updates proposed by Xu, Ma and Kong [13] are essentially the same. Ogasawara and Yabe [10] further extended the convergence results obtained by Yabe and Yamaki [19] to specially structured problems.

At almost the same time and independently, Sheng and Zou [12] studied factorized versions of the structured quasi-Newton methods. The methods are based on the following affine model of r:

$$r(x_k + d) \approx r_k + (J_k + W_k)d.$$

A search direction d_k is given by solving the *linear* least squares problem

$$\text{minimize } \frac{1}{2}\|r_k + (J_k + W_k)d\|^2, \quad d \in \mathbf{R}^n. \tag{1.5}$$

Note that when W_k vanishes, this coincides with the well-known Gauss-Newton model. The solution d_k satisfies the normal equation of (1.5),

$$(J_k + W_k)^\top (J_k + W_k)d_k = -(J_k + W_k)^\top r_k. \tag{1.6}$$

Because of the presence of the vector $-W_k^\top r_k$ on the right-hand side of (1.6), however, the above equation does not correspond to the Newton equation

$$(J_k^\top J_k + S_k)d_k = -J_k^\top r_k.$$

Thus, in order to identify (1.6) with (1.4), Sheng and Zou [12] imposed the additional condition $W_k^\top r_k = 0$ on the matrix W_k. They obtained a BFGS-like update and stated, without a full proof, local and q-superlinear convergence of their method for nonzero residual problems. They proposed a hybrid method which combines their BFGS-like update with the Gauss-Newton update, and presented limited numerical results. Yabe and Takahashi [18] clarified the derivation of Sheng and Zou's BFGS-like update, and reported more detailed computational experiments. They compared the numerical performance among several factorized structured quasi-Newton methods including the Sheng-Zou-BFGS method. Yabe [15] extended Sheng and Zou's BFGS-like update to the Broyden-like family. He also investigated the numerical behavior of a related structured Broyden family of updates obtained from the factorized structured Broyden family (the so-called A-updates; see (3.12) below).

The idea of Sheng and Zou seems interesting for us in that equation (1.4) can be interpreted as the normal equation derived from the linear least squares approximation. Yabe and Takahashi [16] proposed solving equation (1.4) by taking account into only positive definiteness of the coefficient matrix. However, the normal equation (1.6) enables us to reduce the condition number of the coefficient matrix by, for example, using the QR decomposition of the matrix $J_k + W_k$. This is important in determining a search direction d_k by solving the linear equation, and is a significant merit of the Sheng-Zou approach superior to the Yabe-Takahashi approach. Numerical results by Yabe and Takahashi [18] and Yabe [15] suggest that the methods based on the idea of Sheng and Zou are at least competitive to other structured quasi-Newton methods, and the Sheng-Zou approach is more promising than the Yabe-Takahashi approach.

Our concern in this paper is with a theoretical aspect of the Sheng-Zou approach, especially local convergence property of the method with the Sheng-Zou-Broyden family proposed by Yabe [15]. So far, there have been no local convergence results about this method except the result given by Sheng and Zou [12] for their BFGS-like method. It seems, however, that they states local and q-superlinear convergence of their method without giving a complete proof. The purpose of this paper is to give the first proof of local and q-superlinear convergence of the Sheng-Zou-Broyden method, which includes, as a special case, the Sheng-Zou-BFGS method.

The present paper proceeds as follows. In Section 2, we describe the notation and assumptions. In Section 3, we review the standard Broyden, factorized Broyden, structured Broyden and factorized structured Broyden families of updates, one of which is used in the Sheng-Zou-Broyden method analyzed in this paper. In Section 4, we present preliminary results that are unrelated to the methods and generalize a result used in Ogasawara and Yabe [10]. In Section 5, we first show some technical lemmas, and then establish local and q-superlinear convergence of the Sheng-Zou-Broyden method for nonzero residual problems.

2. Notation and Assumptions

Throughout this paper, we denote simply by $\|\cdot\|$ the l_2-norm for vectors or the induced matrix norm. We denote by $\|\cdot\|_F$ the Frobenius norm, i.e., for $A = (a_{ij}) \in \mathbf{R}^{m\times n}$

$$\|A\|_F \equiv [\mathrm{tr}(A^\top A)]^{1/2} = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2\right)^{1/2}.$$

For a given symmetric positive semidefinite matrix $A \in \mathbf{R}^{n\times n}$, $A^{1/2} \in \mathbf{R}^{n\times n}$ denotes any (fixed) symmetric matrix such that $(A^{1/2})^2 = A$. The matrix $A^{-1/2}$ stands for $(A^{1/2})^{-1}$ if A is nonsingular.

We make the following standing assumptions.

(A1) The residual function $r : \mathbf{R}^n \to \mathbf{R}^m$ is twice differentiable in an open convex subset Ω of $\mathbf{R}^n$.

(A2) The nonlinear least squares problem (1.1) has a local minimizer x_*, i.e., there exist a point x_* and a positive constant ε_* such that

$$\|r(x_*)\| \le \|r(x)\| \tag{2.1}$$

for all $x \in \mathcal{D}_* \equiv \{x \in \mathbf{R}^n : \|x - x_*\| < \varepsilon_*\} \subset \Omega$.

(A3) The Jacobian matrix J of r and the Hessian matrix G of f are locally Hölder continuous at x_*, i.e., there exist constants $\xi_J \ge 0$, $\xi_G \ge 0$, $p \in (0, 1]$ and $\varepsilon_c > 0$ such that

$$\|J(x) - J(x_*)\| \le \xi_J \|x - x_*\|^p, \tag{2.2}$$

$$\|G(x) - G(x_*)\| \le \xi_G \|x - x_*\|^p \tag{2.3}$$

for all $x \in \mathcal{D}_c \equiv \{x \in \mathbf{R}^n : \|x - x_*\| < \varepsilon_c\} \subset \Omega$.

(A4) The matrix $G(x_*)$ is positive definite, i.e., there exist constants $\nu_{\max} \ge \nu_{\min} > 0$ such that

$$\nu_{\min}\|u\|^2 \le u^\top G(x_*) u \le \nu_{\max}\|u\|^2 \qquad \text{for all } u \in \mathbf{R}^n.$$

From (2.2) and (2.3) of assumption (A3), it follows that for any $x \in \mathcal{D}_c$

$$\|r(x) - r(x_*) - J(x_*)(x - x_*)\| \le \frac{\xi_J}{p+1}\|x - x_*\|^{p+1}, \tag{2.4}$$

$$\|g(x) - g(x_*) - G(x_*)(x - x_*)\| \le \frac{\xi_G}{p+1}\|x - x_*\|^{p+1}. \tag{2.5}$$

For a proof, see Lemma 4.1.12 in Dennis and Schnabel [7] (see also Exercise 4.4.8). Also, it can be easily seen from assumption (A4) that

$$\|G(x_*)\| \le \nu_{\max} \quad \text{and} \quad \|G(x_*)^{-1}\| \le \frac{1}{\nu_{\min}}. \tag{2.6}$$

We use the notation throughout the paper

$$\sigma(u,v) \equiv \max(\|u-x_*\|^p, \|v-x_*\|^p), \qquad \sigma_k \equiv \sigma(x_k, x_{k+1}),$$

and G_* refers to $G(x_*)$, likewise $J_* = J(x_*)$, $g_k = g(x_k)$, etc.

In what follows, we will often use the following basic relations

$$\begin{aligned}
&\|uv^\top\| = \|uv^\top\|_F = \|u\|\|v\| && \text{for all } u \in \mathbf{R}^m,\ v \in \mathbf{R}^n, \\
&\|A^\top A\| = \|A\|^2, \quad \|A\| \le \|A\|_F && \text{for all } A \in \mathbf{R}^{m\times n}, \\
&\|AB\|_F \le \min\{\|A\|\|B\|_F, \|A\|_F\|B\|\} && \text{for all } A \in \mathbf{R}^{l\times m},\ B \in \mathbf{R}^{m\times n}, && (2.7) \\
&\nu_{\min}\|A\|_* \le \|A\|_F \le \nu_{\max}\|A\|_* && \text{for all } A \in \mathbf{R}^{n\times n}, && (2.8)
\end{aligned}$$

where $\|\cdot\|_*$ denotes the Frobenius norm weighted by G_*, i.e., for $A \in \mathbf{R}^{n\times n}$

$$\|A\|_* \equiv \|G_*^{-1/2} A G_*^{-1/2}\|_F.$$

Inequality (2.7) follows immediately from

$$\|AB\|_F^2 = \sum_{i=1}^n \|Ab_i\|^2 \le \sum_{i=1}^n \|A\|^2\|b_i\|^2 = \|A\|^2\|B\|_F^2\ ,$$

where $b_i \in \mathbf{R}^m$ denotes the ith column vector of B. Relation (2.8) follows from (2.6) and the fact that

$$\begin{aligned}
\|A\|_* &= \|G_*^{-1/2}AG_*^{-1/2}\|_F \le \|G_*^{-1/2}\|^2\|A\|_F = \|G_*^{-1}\|\|A\|_F\ , \\
\|A\|_F &= \|G_*^{1/2}(G_*^{-1/2}AG_*^{-1/2})G_*^{1/2}\|_F \le \|G_*^{1/2}\|^2\|G_*^{-1/2}AG_*^{-1/2}\|_F = \|G_*\|\|A\|_*\ .
\end{aligned}$$

In the rest of this paper, to simplify notation, we drop subscript k and replace "$k+1$" by "+" if not necessary.

3. Factorized Structured Broyden Families

In this section, we review factorized and/or structured versions of the Broyden family of updates. We begin by stating the standard (i.e., unstructured) Broyden family for general unconstrained minimization. Then we describe its factorized form. The corresponding structured Broyden families for nonlinear least squares problems can be expressed by using these families.

Broyden's one-parameter family of updates is defined by

$$B_+^\phi = B + \Delta\text{Broy}(s, y, B, \phi), \tag{3.1}$$

where

$$\Delta\text{Broy}(s,y,B,\phi) \equiv \frac{yy^\top}{s^\top y} - \frac{Bss^\top B}{s^\top Bs} + \phi(s^\top Bs)vv^\top, \tag{3.2}$$

ϕ is a scalar parameter and v is the vector

$$v = v(s, y, B) \equiv \frac{y}{s^\top y} - \frac{Bs}{s^\top Bs}, \tag{3.3}$$

and the vectors s and y are usually defined by

$$s = x_+ - x, \quad y = g_+ - g. \tag{3.4}$$

In (3.2), $s^\top y$ and $s^\top Bs$ are implicitly assumed to be nonzero. Most commonly, it is assumed that $s^\top y > 0$, and that B is symmetric positive definite. Under these assumptions (together with a lower bound assumption on ϕ), it is known that B_+^ϕ is also symmetric positive definite. The updates B_+^ϕ can be written as

$$B_+^\phi = (1-\phi)B_+^0 + \phi B_+^1.$$

The updates B_+^ϕ for $\phi \in [0,1]$ are called the *convex* class, and the extremes B_+^0 and B_+^1 are the well-known BFGS and DFP updates.

For later use, we define here the DFP update separately as follows:

$$B_+^{\mathrm{DFP}} \equiv B + \Delta\mathrm{DFP}(s, y, B),$$

where

$$\Delta\mathrm{DFP}(s, y, B) \equiv \frac{(y - Bs)y^\top + y(y - Bs)^\top}{s^\top y} - \frac{s^\top(y - Bs)}{(s^\top y)^2} yy^\top. \tag{3.5}$$

Needless to say, $\Delta\mathrm{DFP}(s, y, B) = \Delta\mathrm{Broy}(s, y, B, 1)$, and $B_+^{\mathrm{DFP}} = B_+^1$. Since $B_+^1 - B_+^0 = (s^\top Bs)vv^\top$, the Broyden family B_+^ϕ can be rewritten by using the DFP update as

$$\begin{aligned} B_+^\phi &= B_+^1 + (\phi - 1)(B_+^1 - B_+^0) \\ &= B_+^{\mathrm{DFP}} + (\phi - 1)\Delta B(s, y, B), \end{aligned} \tag{3.6}$$

where

$$\Delta B(s, y, B) \equiv (s^\top Bs)\Big(\frac{y}{s^\top y} - \frac{Bs}{s^\top Bs}\Big)\Big(\frac{y}{s^\top y} - \frac{Bs}{s^\top Bs}\Big)^\top. \tag{3.7}$$

We will use this form later because it is convenient for showing our local convergence properties.

Yabe [14] constructed a *factorized* form of the standard Broyden family (3.1)–(3.4). In the companion paper, Yamaki and Yabe [20] further studied and clarified these updates. Actually, Yabe [14] and Yamaki and Yabe [20] dealt with *inverse* updates, and obtained the factorized form of the *sized* Broyden family as a special case of more general results. Here the term *size* is used in the sense of Oren and Luenberger [11], who used the term *scale* instead. In the present paper, however, we only deal with a factorized form of the *unsized* Broyden family of *direct* updates. Its formal expression is presented in the following:

$$N_+ = N + \Delta\mathrm{FacBroy}(s, y, N, \phi), \tag{3.8}$$

where

$$\Delta\mathrm{FacBroy}(s, y, N, \phi) \equiv (1 - \sqrt{\phi})\frac{Ns}{s^\top Bs}(\sqrt{\lambda}y - Bs)^\top + \sqrt{\phi}N(\sqrt{\lambda}B^{-1}y - s)\frac{y^\top}{s^\top y}, \tag{3.9}$$

and

$$B = N^\top N, \quad \phi \geq 0, \quad \lambda = \lambda(s, y, N, \phi) \equiv \Big((1-\phi)\frac{s^\top y}{s^\top Bs} + \phi\frac{y^\top B^{-1}y}{s^\top y}\Big)^{-1} \tag{3.10}$$

with s and y given by (3.4). Here we assume that $s^\top y > 0$, and that the matrix N is an $m \times n$ rectangular matrix with full column rank. Note that the latter assumption ensures that B is symmetric positive definite, and particularly $s^\top Bs > 0$. (Actually, in Yabe [14] and Yamaki and Yabe [20], N was assumed to be a nonsingular square matrix of order n.)

Yabe [14] (see also Yamaki and Yabe [20]) proved that

$$B_+ = N_+^\top N_+ = B + \Delta\mathrm{Broy}(s, y, B, \phi). \tag{3.11}$$

Thus, we call the family (3.8)–(3.10) a *factorized Broyden family* of N. From (3.11) and the assumption $s^\top y > 0$, in the same way as the ordinary Broyden updates, it follows that B_+ is also symmetric positive definite, which means N_+ is also of full column rank.

Although we do not discuss a structured version of the Broyden family in this paper, we mention it according to the structure principle given by Dennis, Martínez and Tapia [5]. Replacing B_+, B and y in (3.1) by, respectively,

$$B_+ = J_+^\top J_+ + A_+, \quad B^\sharp = J_+^\top J_+ + A \quad \text{and} \quad y^\sharp = J_+^\top J_+ s + y^\flat,$$

where $y^\flat$ is an approximation to $S_+ s$, we have

$$B_+ = B^\sharp + \Delta\mathrm{Broy}(s, y^\sharp, B^\sharp, \phi).$$

This implies the update

$$A_+ = A + \Delta\mathrm{Broy}(s, J_+^\top J_+ s + y^\flat, J_+^\top J_+ + A, \phi). \tag{3.12}$$

We call this the *structured Broyden family* of A, or shortly, the A-update. A typical choice of $y^\flat$ is the vector

$$y^\flat = y^{\mathrm{DB}\flat} \equiv (J_+ - J)^\top r_+, \tag{3.13}$$

which was proposed by Dennis [3] and, independently, by Bartholomew-Biggs [2]. This vector is most preferred to others and is widely used because several authors reported that this choice improved numerical performance (see, for example, [1] or [2]).

The *structured* quasi-Newton updates differ from the ordinary *unstructured* quasi-Newton updates in that they use the intermediate information available. Specifically, the next B_+ is generated by using $B^\sharp$, not using the previous B. Therefore, whereas the ordinary quasi-Newton methods need only $O(n^2)$ arithmetic operations at each iteration step, the structured quasi-Newton methods generally require $O(mn^2)$ arithmetic operations. Nevertheless, the structured quasi-Newton methods are widely recognized to be efficient by the popularity and success of the NL2SOL code of Dennis, Gay and Welsh [4]. It is probably most important that since the structured quasi-Newton updates use the new exact value of the Jacobian matrix, those may have more better information about the Hessian than the ordinary quasi-Newton updates approximating the overall Hessian matrix.

We next consider two structured versions of the factorized Broyden family. These families, of course, can also be viewed as factorized forms of the structured Broyden family. (See Figure 1.)

Before stating the families, we introduce some notations to describe those. We first define the matrices

$$Q \equiv \frac{r_+ r_+^\top}{\|r_+\|^2} \quad \text{and} \quad P \equiv I - Q = I - \frac{r_+ r_+^\top}{\|r_+\|^2}. \tag{3.14}$$

Clearly, P and Q are $m \times m$ orthogonal projection matrices, and so $P^2 = P$, $P^\top = P$, $Q^2 = Q$, $Q^\top = Q$, $PQ = 0$, $\|P\| = \|Q\| = \|Q\|_F = 1$. Furthermore, we define the following matrices and vectors:

$$\begin{aligned}
N_+ &= J_+ + W_+, \quad B_+ = N_+^\top N_+, \\
N^\sharp &\equiv J_+ + W, \quad B^\sharp \equiv N^{\sharp\top} N^\sharp, \\
N^\natural &\equiv J_+ + PW, \quad B^\natural \equiv N^{\natural\top} N^\natural, && (3.15) \\
N^{P\natural} &\equiv PN^\natural = P(J_+ + PW), \quad B^{P\natural} \equiv N^{P\natural\top} N^{P\natural} = N^{\natural\top} P N^\natural, && (3.16) \\
N^{Q\natural} &\equiv QN^\natural = Q(J_+ + PW), \quad B^{Q\natural} \equiv N^{Q\natural\top} N^{Q\natural} = N^{\natural\top} Q N^\natural, && (3.17) \\
y^\sharp &\equiv J_+^\top J_+ s + y^\flat \quad \text{and} \quad y^{P\natural} \equiv J_+^\top P J_+ s + y^\flat, && (3.18)
\end{aligned}$$

where again $y^\flat$ is an approximation to $S_+ s$. We can readily verify that

$$\begin{aligned}
N^\natural &= N^{P\natural} + N^{Q\natural}, \quad B^\natural = B^{P\natural} + B^{Q\natural}, && (3.19) \\
N^{P\natural} &= P(J_+ + W) = PN^\sharp, \quad B^{P\natural} = (PN^\sharp)^\top P N^\sharp = N^{\sharp\top} P N^\sharp, && (3.20) \\
N^{Q\natural} &= QJ_+, \quad B^{Q\natural} = (QJ_+)^\top Q J_+ = J_+^\top Q J_+, && (3.21) \\
y^{P\natural} &= J_+^\top J_+ s + y^\flat - J_+^\top Q J_+ s = y^\sharp - B^{Q\natural} s. && (3.22)
\end{aligned}$$

By replacing N and y in (3.8) by, respectively, $N^\sharp$ and $y^\sharp$, we obtain the first structured version of the factorized Broyden family as follows:

$$N_+ = N^\sharp + \Delta\mathrm{FacBroy}(s, y^\sharp, N^\sharp, \phi), \tag{3.23}$$

which yields the update

$$W_+ = W + \Delta\mathrm{FacBroy}(s, J_+^\top J_+ s + y^\flat, J_+ + W, \phi). \tag{3.24}$$

Similarly, by replacing two N's (the first term and the third argument of ΔFacBroy on the right-hand side) and y in (3.8) by, respectively, $N^\natural$, $N^{P\natural}$ and $y^{P\natural}$, we obtain the second structured version of the factorized Broyden family as follows:

$$N_+ = N^\natural + \Delta\mathrm{FacBroy}(s, y^{P\natural}, N^{P\natural}, \phi), \tag{3.25}$$

which gives the update

$$W_+ = PW + \Delta\mathrm{FacBroy}(s, J_+^\top P J_+ s + y^\flat, P(J_+ + W), \phi). \tag{3.26}$$

We call (3.24) and (3.26) the *Yabe-Yamaki-Broyden* (*YY-Broyden*) and *Sheng-Zou-Broyden* (*SZ-Broyden*) *families* of W, respectively. The family (3.24) was proposed by Yabe and Yamaki [19] who called a factorized Broyden-like family, while the family (3.26) was proposed by Yabe [15] who derived as a generalization of a BFGS-like update originally introduced by Sheng and Zou [12]. Sheng and Zou's BFGS-like update is obtained as a special case of the Sheng-Zou-Broyden updates only by setting $\phi = 0$. Yabe [15] mainly dealt with the convex class of the family. Incidentally, Yabe and Yamaki [19], Yabe [15], and Sheng and Zou [12] dealt with, as a choice of $y^\flat$, only the case of $y^\flat = y^{\mathrm{DB}\flat}$, i.e., the Dennis-Biggs vector (3.13).

We note that in (3.23) we assume that $s^\top y^\sharp > 0$, and that $N^\sharp$ is a rectangular matrix with full column rank. Similarly, in (3.25) we assume that $s^\top y^{P\natural} > 0$, and that $N^{P\natural}$ is a rectangular matrix with full column rank. It should be noticed that in (3.23) the first term and the third argument of the second term on the right-hand side are both the same $N^\sharp$, while in (3.25) those are distinct: $N^\natural$ and $N^{P\natural}$.

By (3.11), we have

$$B_+ = B^\sharp + \Delta\mathrm{Broy}(s, y^\sharp, B^\sharp, \phi), \tag{3.27}$$

which ensures that N_+ is also of full column rank as before. On the other hand, Yabe [15] showed that

$$B_+ = B^\natural + \Delta\mathrm{Broy}(s, y^{P\natural}, B^{P\natural}, \phi). \tag{3.28}$$

Again we note here that in (3.27) the first term and the third argument of the second term on the right-hand side are both the same $B^\sharp$, while in (3.28) those are distinct: $B^\natural$ and $B^{P\natural}$. Therefore, unlike the former, that N_+ is of full column rank is not a direct consequence of the form (3.28). However, that is shown as follows. By using (3.19), the update (3.28) can be written as

$$B_+ = B_+^P + B^{Q\natural}, \tag{3.29}$$

where

$$B_+^P = B^{P\natural} + \Delta\mathrm{Broy}(s, y^{P\natural}, B^{P\natural}, \phi). \tag{3.30}$$

That the matrix B_+^P is symmetric positive definite is a direct consequence of the assumption $s^\top y^{P\natural} > 0$ and the fact that $B^{P\natural}$ is symmetric positive definite because of the assumption $N^{P\natural}$ having full column rank. Since, by definition, $B^{Q\natural}$ is symmetric positive semidefinite, it follows that B_+ is also symmetric positive definite. Thus, N_+ has full column rank.

Finally, we note that if we formally set $Q = 0$ or, equivalently, $P = I$ in the definitions (3.15)–(3.18), then we have

$$N^{P\natural} = N^\natural = N^\sharp, \quad B^{P\natural} = B^\natural = B^\sharp, \quad N^{Q\natural} = 0, \quad B^{Q\natural} = 0, \quad y^{P\natural} = y^\sharp.$$

This means that when we do not impose the orthogonality condition $W_+^\top r_+ = 0$ on the matrix W_+ for the Sheng-Zou-Broyden family, the family (3.26) reduces to the Yabe-Yamaki-Broyden family (3.24). We summarize the above explanation so far in Figure 1.

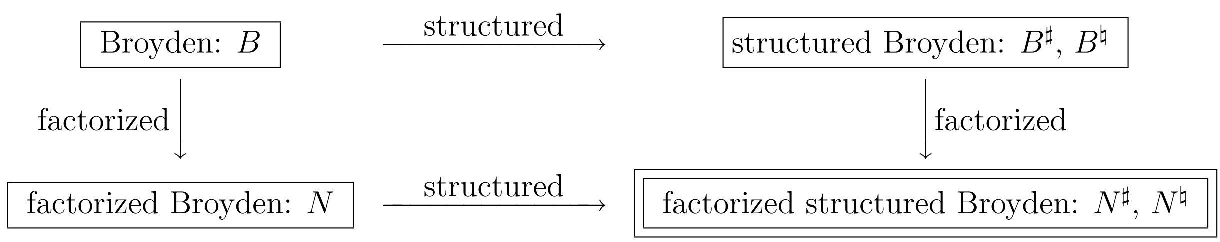

Figure 1: Factorized/structured Broyden families

We are ready to present the following algorithm.

Sheng-Zou-Broyden (SZ-Broyden) Method.

Step 0. Choose an initial point $x_0 \in \mathbf{R}^n$ and an initial matrix $W_0 \in \mathbf{R}^{m\times n}$ (usually $W_0 := 0$), and set $k := 0$.

Step 1. Given $x_k \in \mathbf{R}^n$ and $W_k \in \mathbf{R}^{m\times n}$, solve the linear system of equations for s_k,

$$(J_k + W_k)^\top (J_k + W_k) s_k = -(J_k + W_k)^\top r_k,$$

or, equivalently, the linear least squares problem

$$\text{minimize } \frac{1}{2}\|r_k + (J_k + W_k)s_k\|^2, \quad s_k \in \mathbf{R}^n.$$

Step 2. Set $x_{k+1} := x_k + s_k$.
Step 3. Compute P_k, $y_k^{P\natural}$, $N_k^{P\natural}$ and W_{k+1} by

$$\begin{aligned} P_k &= I - \frac{r_{k+1}r_{k+1}^\top}{\|r_{k+1}\|^2}, \\ y_k^{P\natural} &= J_{k+1}^\top P_k J_{k+1} s_k + y_k^\flat, \\ N_k^{P\natural} &= P_k(J_{k+1} + P_k W_k), \\ W_{k+1} &= P_k W_k + \Delta\mathrm{FacBroy}(s_k, y_k^{P\natural}, N_k^{P\natural}, \phi_k), \end{aligned}$$

where $y_k^\flat$ is a suitably defined approximation to $S_{k+1}s_k$, and $\phi_k \geq 0$.
Step 4. Set $k := k+1$, and return to Step 1.

In this paper, we deal with the factorized structured Broyden family from the *nonnegatively bounded* class, i.e., $\phi \in [0, \phi_U]$ for some constant ϕ_U.

Similar to (3.6), the matrix (3.30) can be rewritten as

$$B_+^P = B_+^{\mathrm{DFP}\natural} + (\phi - 1)\Delta B(s, y^{P\natural}, B^{P\natural}), \tag{3.31}$$

where

$$B_+^{\mathrm{DFP}\natural} \equiv B^{P\natural} + \Delta\mathrm{DFP}(s, y^{P\natural}, B^{P\natural}),$$

ΔDFP and ΔB are defined by (3.5) and (3.7), respectively. Substituting (3.31) into (3.29), we can write the structured Broyden family (3.28) as

$$B_+ = \left\{ B_+^{\mathrm{DFP}\natural} + (\phi - 1)\Delta B(s, y^{P\natural}, B^{P\natural}) \right\} + B^{Q\natural}. \tag{3.32}$$

This form is convenient for our convergence analysis.

4. Preliminaries

The following inequality is very simple, but is effectively used in Proposition 2.

Proposition 1 *For any $E \in \mathbf{R}^{m\times n}$, it holds*

$$\|E\|_F \leq \sqrt{\mathrm{rank}(E)}\|E\|. \tag{4.1}$$

Proof. We may assume without loss of generality that $m \geq n$. Let $\lambda_1 \geq \cdots \geq \lambda_r > 0$ denote r nonzero eigenvalues of the matrix $E^\top E$, so that the rest are $n-r$ zero eigenvalues. Here $r = \mathrm{rank}(E^\top E) = \mathrm{rank}(E)$ and it is possibly zero. Then we have

$$\|E\|_F^2 = \mathrm{tr}(E^\top E) = \sum_{i=1}^r \lambda_i \leq r\lambda_1 = \mathrm{rank}(E)\|E\|^2,$$

which is the desired result. □

The following proposition is a generalization of Lemma 3 in Ogasawara and Yabe [10].

Proposition 2 *Let $l \geq m$, $A \in \mathbf{R}^{l\times n}$ and B in $\mathbf{R}^{m\times n}$. Define*

$$\begin{aligned} \mathcal{A} &\equiv \{\tilde{A} \in \mathbf{R}^{l\times n} : \textit{Rows of } \tilde{A} \textit{ consist of a permutation of rows of } A\}, \\ \mathcal{B} &\equiv \{\tilde{B} \in \mathbf{R}^{l\times n} : \textit{Rows of } \tilde{B} \textit{ consist of a permutation of rows of } \begin{bmatrix} B \\ 0 \end{bmatrix}, \\ &\qquad \textit{where } 0 \in \mathbf{R}^{(l-m)\times n}\}. \end{aligned}$$

Then

$$\|A^\top A - B^\top B\| \leq \min_{\substack{\tilde{A}\in\mathcal{A}\\ \tilde{B}\in\mathcal{B}}} \{\|\tilde{A}+\tilde{B}\|\|\tilde{A}-\tilde{B}\|\}, \tag{4.2}$$

$$\|A^\top A - B^\top B\|_F \leq \min_{\substack{\tilde{A}\in\mathcal{A}\\ \tilde{B}\in\mathcal{B}}} \{\sqrt{\min\{\mathrm{rank}(\tilde{A}+\tilde{B}), \mathrm{rank}(\tilde{A}-\tilde{B})\}}\|\tilde{A}+\tilde{B}\|\|\tilde{A}-\tilde{B}\|\}. \tag{4.3}$$

Proof. We only give the proof of (4.3). Inequality (4.2) is also proved similarly (actually more easily). Choose any $\tilde{A} \in \mathcal{A}$ and $\tilde{B} \in \mathcal{B}$. Then it is clear that $A^\top A = \tilde{A}^\top \tilde{A}$ and $B^\top B = \tilde{B}^\top \tilde{B}$. Since

$$A^\top A - B^\top B = \tilde{A}^\top \tilde{A} - \tilde{B}^\top \tilde{B} = \frac{1}{2}\Big\{(\tilde{A}+\tilde{B})^\top(\tilde{A}-\tilde{B}) + (\tilde{A}-\tilde{B})^\top(\tilde{A}+\tilde{B})\Big\},$$

we have

$$\begin{aligned} \|A^\top A - B^\top B\|_F &\leq \frac{1}{2}\Big\{\|(\tilde{A}+\tilde{B})^\top(\tilde{A}-\tilde{B})\|_F + \|(\tilde{A}-\tilde{B})^\top(\tilde{A}+\tilde{B})\|_F\Big\} \\ &= \|(\tilde{A}+\tilde{B})^\top(\tilde{A}-\tilde{B})\|_F \\ &\leq \min\{\|\tilde{A}+\tilde{B}\|_F\|\tilde{A}-\tilde{B}\|, \|\tilde{A}+\tilde{B}\|\|\tilde{A}-\tilde{B}\|_F\} \\ &\leq \min\{\sqrt{\mathrm{rank}(\tilde{A}+\tilde{B})}\|\tilde{A}+\tilde{B}\| \cdot \|\tilde{A}-\tilde{B}\|, \\ &\qquad \|\tilde{A}+\tilde{B}\| \cdot \sqrt{\mathrm{rank}(\tilde{A}-\tilde{B})}\|\tilde{A}-\tilde{B}\|\} \\ &= \sqrt{\min\{\mathrm{rank}(\tilde{A}+\tilde{B}), \mathrm{rank}(\tilde{A}-\tilde{B})\}}\|\tilde{A}+\tilde{B}\|\|\tilde{A}-\tilde{B}\|, \end{aligned}$$

where the second inequality follows from (2.7), and the third inequality follows from (4.1). □

Corollary 1 (Factorization inequalities) *For any matrices A and B in $\mathbf{R}^{m\times n}$,*

$$\|A^\top A - B^\top B\| \leq \|A+B\|\|A-B\|, \tag{4.4}$$

$$\|A^\top A - B^\top B\|_F \leq \sqrt{\min(m,n)}\|A+B\|\|A-B\|. \tag{4.5}$$

In particular, for any vectors a and b in $\mathbf{R}^n$,

$$\|aa^\top - bb^\top\|_F \leq \|a+b\|\|a-b\|. \tag{4.6}$$

Proof. Since clearly $A \in \mathcal{A}$ and $B \in \mathcal{B}$ (note that in this case $l = m$), the result follows directly from Proposition 2 and $\mathrm{rank}(A \pm B) \leq \min(m,n)$. □

Remark 1 Direct application of the inequality $\|E\|_F \leq \sqrt{\min(m,n)}\|E\|$ for $E \in \mathbf{R}^{m\times n}$ to estimating the Frobenius norm yields together with (4.4)

$$\begin{aligned} \|A^\top A - B^\top B\|_F &\leq \sqrt{n}\|A^\top A - B^\top B\| \\ &\leq \sqrt{n}\|A+B\|\|A-B\|, \\ \|aa^\top - bb^\top\|_F &\leq \sqrt{n}\|a+b\|\|a-b\|, \end{aligned}$$

which are worse than (4.5) and (4.6). Instead, if we use $\|E\|_F \le \sqrt{\mathrm{rank}(E)}\|E\|$, then we have better inequalities whose coefficients are respectively $\sqrt{\min(2m,n)}$ and $\sqrt{2}$, although these are a little yet worse than (4.5) and (4.6). Proposition 2 gives smaller bounds than Corollary 1 as the following example shows. Let $A = \begin{bmatrix} C \\ D \end{bmatrix}$, $B = \begin{bmatrix} D \\ C \end{bmatrix}$. The matrices C and D may be of different sizes. Then inequality (4.3) gives the exact bound zero, whereas (4.5) does not. Corollary 1 (Lemma 3 of Ogasawara and Yabe [10]) is sufficient for our purpose in this paper, but the inequalities in Proposition 2 might be useful somewhere else.

5. Local and Q-Superlinear Convergence of the SZ-Broyden Method

In this section, we show local and q-superlinear convergence of the Sheng-Zou-Broyden method for *nonzero* residual problems. To this end, we need some technical lemmas.

For notational convenience, let us set

$$\widehat{B}^{P\natural} := G_*^{-1/2} B^{P\natural} G_*^{-1/2}, \quad \widehat{y}^{P\natural} := G_*^{-1/2} y^{P\natural} \quad \text{and} \quad \widehat{s} := G_*^{1/2} s.$$

Lemma 1 *There holds*

$$\|QJ_+\|_F = \|QJ_+\| = \|B^{Q\natural}\|^{1/2} = \|B^{Q\natural}\|_F^{1/2} = \frac{\|J_+^\top r_+\|}{\|r_+\|}. \tag{5.1}$$

Suppose that assumptions (A1)–(A3) *hold for a nonzero residual problem, i.e.,* $\|r_*\| > 0$. *Then, there exist positive constants* $\tilde{\zeta}_g$ *and* ζ_g *such that*

$$\|QJ_+\| = \|B^{Q\natural}\|^{1/2} \le \tilde{\zeta}_g \|x_+ - x_*\| \le \zeta_g \sigma(x, x_+) \tag{5.2}$$

for all $x, x_+ \in \mathcal{D}_* \cap \mathcal{D}_c$.

Proof. By the definition (3.14) of Q, we have

$$\|QJ_+\|_F = \left\| \frac{r_+ r_+^\top}{\|r_+\|^2} J_+ \right\|_F = \frac{\|J_+^\top r_+\|}{\|r_+\|} = \|QJ_+\|.$$

Similarly, by (3.21), we get

$$\|B^{Q\natural}\|_F = \left\| J_+^\top \frac{r_+ r_+^\top}{\|r_+\|^2} J_+ \right\|_F = \frac{\|J_+^\top r_+\|^2}{\|r_+\|^2} = \|QJ_+\|^2 = \|B^{Q\natural}\|.$$

Assume that assumptions (A1)–(A3) hold with $\|r_*\| > 0$. Set $\varepsilon_0 := \min(\varepsilon_*, \varepsilon_c)$ and let $x, x_+ \in \mathcal{D}_* \cap \mathcal{D}_c$. Using (2.1), (2.5) and (2.6), we have

$$\begin{aligned}
\|QJ_+\| &= \frac{\|J_+^\top r_+\|}{\|r_+\|} = \frac{\|g_+\|}{\|r_+\|} = \frac{\|g_+ - g_*\|}{\|r_+\|} \\
&\le \frac{\|g_+ - g_* - G_*(x_+ - x_*)\| + \|G_*(x_+ - x_*)\|}{\|r_*\|} \\
&\le \frac{1}{\|r_*\|}\left(\frac{\xi_G}{p+1}\|x_+ - x_*\|^{p+1} + \|G_*\|\|x_+ - x_*\| \right) \\
&\le \frac{1}{\|r_*\|}\left(\frac{\xi_G}{p+1}\varepsilon_0^p + \nu_{\max} \right)\|x_+ - x_*\| \\
&= \tilde{\zeta}_g \|x_+ - x_*\| \\
&\le \tilde{\zeta}_g \varepsilon_0^{1-p} \|x_+ - x_*\|^p \\
&\le \zeta_g \sigma(x, x_+)
\end{aligned}$$

with $\tilde{\zeta}_g := \left(\nu_{\max} + \frac{\xi_G \varepsilon_0^p}{p+1}\right)/\|r_*\| > 0$ and $\zeta_g := \tilde{\zeta}_g \varepsilon_0^{1-p} = \left(\nu_{\max}\varepsilon_0^{1-p} + \frac{\xi_G \varepsilon_0}{p+1}\right)/\|r_*\| > 0$. □

Lemma 2 *There holds*

$$\|QW\|_F = \|QW\| = \frac{\|W^\top r_+\|}{\|r_+\|}.$$

Suppose that assumptions (A1)–(A3) *hold for a nonzero residual problem, and that W satisfies $W^\top r = 0$. Then, there exists a positive constant ζ_r such that*

$$\|QW\| \le \zeta_r \|W\| \sigma(x, x_+) \quad \textit{for all } x, x_+ \in \mathcal{D}_* \cap \mathcal{D}_c. \tag{5.3}$$

Proof. Similar to (5.1) of Lemma 1, we have

$$\|QW\|_F = \left\| \frac{r_+ r_+^\top}{\|r_+\|^2} W \right\|_F = \frac{\|W^\top r_+\|}{\|r_+\|} = \|QW\|.$$

Assume that assumptions (A1)–(A3) hold with $\|r_*\| > 0$ and $W^\top r = 0$. Set again $\varepsilon_0 := \min(\varepsilon_*, \varepsilon_c)$ and let $x, x_+ \in \mathcal{D}_* \cap \mathcal{D}_c$. Using $W^\top r = 0$, (2.1) and (2.4), we have

$$\begin{aligned}
\|QW\| &= \frac{\|W^\top r_+\|}{\|r_+\|} = \frac{\|W^\top (r_+ - r)\|}{\|r_+\|} \\
&\le \frac{\|W\|}{\|r_*\|}\Big(\|r_+ - r_* - J_*(x_+ - x_*)\| + \| - r + r_* + J_*(x - x_*)\| + \|J_*(x_+ - x)\|\Big) \\
&\le \frac{\|W\|}{\|r_*\|}\Big(\frac{\xi_J}{p+1}\{\|x_+ - x_*\|^{p+1} + \|x - x_*\|^{p+1}\} + \|J_*\|\{\|x_+ - x_*\| + \|x - x_*\|\}\Big) \\
&\le \frac{\|W\|}{\|r_*\|}\Big(\frac{\xi_J \varepsilon_0}{p+1} + \|J_*\|\varepsilon_0^{1-p}\Big)(\|x_+ - x_*\|^p + \|x - x_*\|^p) \\
&\le \zeta_r \|W\| \sigma(x, x_+)
\end{aligned}$$

with $\zeta_r := 2\left(\|J_*\|\varepsilon_0^{1-p} + \frac{\xi_J \varepsilon_0}{p+1}\right)/\|r_*\| > 0$. □

We note that the following lemma is dependent only on how to take $y^\sharp$ or $y^{P\natural}$, and independent of an algorithm, particularly how to choose a matrix W.

Lemma 3 *Suppose that assumptions* (A1) *and* (A3) *hold. Assume that there exist constants $\zeta_\flat \ge 0$ and $\varepsilon_\flat > 0$ such that $y^\flat$ satisfies*

$$\|y^\flat - S_* s\| \le \zeta_\flat \sigma(x, x_+)\|s\| \qquad \textit{for all } x, x_+ \in \mathcal{D}_\flat, \tag{5.4}$$

where $s := x_+ - x$, $\mathcal{D}_\flat \equiv \{x \in \mathbf{R}^n : \|x - x_*\| < \varepsilon_\flat\}$. Then, there exists a nonnegative constant $\zeta_\sharp$ such that $y^\sharp \equiv J_+^\top J_+ s + y^\flat$ satisfies

$$\|y^\sharp - G_* s\| \le \zeta_\sharp \sigma(x, x_+)\|s\| \qquad \textit{for all } x, x_+ \in \mathcal{D}_c \cap \mathcal{D}_\flat.$$

If, in addition, assumptions (A2) *and* (A4) *hold for a nonzero residual problem, then there exist constants $\widehat{\zeta}_\natural \ge 0$, $\gamma \ge 1$ and $\varepsilon_\natural > 0$ such that*

$$\begin{aligned}
\|\widehat{y}^{P\natural} - \widehat{s}\| &\le \widehat{\zeta}_\natural \sigma(x, x_+)\|\widehat{s}\|, &(5.5)\\
0 < \frac{\|\widehat{s}\|\|\widehat{y}^{P\natural}\|}{\widehat{s}^\top \widehat{y}^{P\natural}} &\le \gamma, &(5.6)
\end{aligned}$$

for all distinct $x, x_+ \in \mathcal{D}_\natural \equiv \{x \in \mathbf{R}^n : \|x - x_*\| < \varepsilon_\natural\} \subset \mathcal{D}_* \cap \mathcal{D}_c \cap \mathcal{D}_\flat$.

Proof. The first half is a special case of Lemma 4 in Ogasawara and Yabe [10]. (Take $E := J$, $\mathcal{D}_1 := \mathcal{D}_c$, $\mathcal{D} := \mathcal{D}_\flat$, $\bar{x} := x_+$ in Lemma 4 of Ogasawara and Yabe [10].)

To show the second half, we note that the deviation $B^{Q\natural}s$ from $y^\sharp$ in $y^{P\natural}$ can be regarded as the deviation from $y^\flat$. Specifically, from the definitions (3.18) of $y^\sharp$ and $y^{P\natural}$ (see also (3.22)), we can write $y^{P\natural}$ as $y^{P\natural} = y^\sharp - B^{Q\natural}s = J_+^\top J_+ s + y^\flat - B^{Q\natural}s = J_+^\top J_+ s + \tilde{y}^\flat$, where we have set $\tilde{y}^\flat := y^\flat - B^{Q\natural}s$. Set $\varepsilon_1 := \min(\varepsilon_*, \varepsilon_c, \varepsilon_\flat)$ and let x, $x_+ \in \mathcal{D}_* \cap \mathcal{D}_c \cap \mathcal{D}_\flat$, $\|r_*\| > 0$. Since $y^\flat$ satisfies (5.4), and by (5.2) of Lemma 1, $\tilde{y}^\flat$ satisfies that

$$\begin{aligned} \|\tilde{y}^\flat - S_* s\| &\leq \|y^\flat - S_* s\| + \|B^{Q\natural}\|\|s\| \\ &\leq \zeta_\flat \sigma(x, x_+)\|s\| + \zeta_g^2 \sigma(x, x_+)^2 \|s\| \\ &\leq (\zeta_\flat + \zeta_g^2 \varepsilon_1^p)\sigma(x, x_+)\|s\| \\ &= \tilde{\zeta}_\flat \sigma(x, x_+)\|s\| \end{aligned}$$

with $\tilde{\zeta}_\flat := \zeta_\flat + \zeta_g^2 \varepsilon_1^p$. Again we can apply Lemma 4 of Ogasawara and Yabe [10]. Replacing E, $\mathcal{D}_1$, $y^\flat$, $\zeta^\flat$, $\mathcal{D}$, $y^\sharp$ and $\bar{x}$ in Lemma 4 of Ogasawara and Yabe [10], by, respectively, J, $\mathcal{D}_c$, $\tilde{y}^\flat$, $\tilde{\zeta}_\flat$, $\mathcal{D}_* \cap \mathcal{D}_c \cap \mathcal{D}_\flat$, $y^{P\natural}$ and x_+, we have the desired result. □

The following two lemmas are essentially the same as Lemmas 5 and 6 given in Ogasawara and Yabe [10], so we omit their proofs.

Lemma 4 *Suppose that all the assumptions of Lemma* 3 *hold, and that, for some positive constant* $\delta_\natural$*, there holds*

$$\|B^{P\natural} - G_*\|_* \leq \delta_\natural.$$

Then, for all distinct x*,* $x_+ \in \mathcal{D}_\natural$*,*

$$\|B_+^{\mathrm{DFP}\natural} - G_*\|_* \leq [1 + \omega_1 \sigma(x, x_+)]\|B^{P\natural} - G_*\|_* + \omega_2 \sigma(x, x_+) - \frac{\|(\widehat{B}^{P\natural} - I)\widehat{s}\|^2}{2\delta_\natural \|\widehat{s}\|^2}, \tag{5.7}$$

where $\omega_1 = (\gamma+1)(\gamma+3)\widehat{\zeta}_\natural \geq 0$*,* $\omega_2 = \gamma(\gamma+2)\widehat{\zeta}_\natural \geq 0$*, and* $\mathcal{D}_\natural$*,* $\widehat{\zeta}_\natural$*,* γ *are given in Lemma* 3*.*

Lemma 5 *Suppose that all the assumptions of Lemma* 4 *hold with* $\delta_\natural < 1$*. Let* ΔB *be defined as in* (3.7)*. Then, for all distinct* x*,* $x_+ \in \mathcal{D}_\natural$*,*

$$\|\Delta B(s, y^{P\natural}, B^{P\natural})\|_* \leq \tau_1(\delta_\natural)\left[\tau_2(\delta_\natural)\sigma(x, x_+) + \frac{\|(\widehat{B}^{P\natural} - I)\widehat{s}\|^2}{\|\widehat{s}\|^2}\right] \tag{5.8}$$

where $\tau_1(\delta_\natural) \equiv \dfrac{(\gamma+1)^2}{1-\delta_\natural} > 0$*,* $\tau_2(\delta_\natural) \equiv (1 + 2\delta_\natural)\widehat{\zeta}_\natural \geq 0$*, and* $\mathcal{D}_\natural$*,* $\widehat{\zeta}_\natural$*,* γ *are given in Lemma* 3*.*

The following lemma is a slight modification of Lemma 7 in Ogasawara and Yabe [10].

Lemma 6 *Suppose that all the assumptions of Lemma* 5 *hold, and that there exists some nonnegative constant* K *independent of* x*,* x_+ *such that*

$$\|B^{P\natural} - B\|_* \leq K\sigma(x, x_+) \qquad \text{for some distinct } x, x_+ \in \mathcal{D}_\natural. \tag{5.9}$$

Let B_+ *be a structured Broyden update derived from the Sheng-Zou-Broyden family with* $\phi \in [0, \phi_U]$ *for some constant* $\phi_U \geq 0$*. Let* $\phi' = \max(1, |\phi_U - 1|)$*. If* $\delta_\natural \leq 0.9\bar{\delta}$*, where* $\bar{\delta} = 1/\{1 + 2\phi'(\gamma+1)^2\}$*, then*

$$\|B_+ - G_*\|_* \leq [1 + \mu_1 \sigma(x, x_+)]\|B - G_*\|_* + \mu_2 \sigma(x, x_+) - \mu_3 \frac{\|(\widehat{B}^{P\natural} - I)\widehat{s}\|^2}{\|\widehat{s}\|^2},$$

where $\mu_1 = \omega_1 \geq 0$, $\mu_2 = (1+\omega_1\varepsilon_\natural^p)K + \omega_2 + \zeta_g^2\varepsilon_\natural^p/\nu_{\min} + \phi'\bar{\tau}_1\bar{\tau}_2 \geq 0$, $\mu_3 = 1/(18\bar{\delta}) > 0$, *and* $\mathcal{D}_\natural$, $\varepsilon_\natural$, γ *are given in Lemma* 3, ω_1, ω_2 *in Lemma* 4, *and* $\bar{\tau}_1 = \tau_1(0.9\bar{\delta})$, $\bar{\tau}_2 = \tau_2(0.9\bar{\delta})$, $\tau_1(\cdot)$, $\tau_2(\cdot)$ *in Lemma* 5.

Proof. Let $x, x_+ \in \mathcal{D}_\natural$, $x \neq x_+$. For simplicity of notation, let $\sigma = \sigma(x, x_+)$. Note from the definition that $|\phi - 1| \leq \phi'$. We want to estimate $B_+ - G_*$, and we have already estimated $B_+^{\mathrm{DFP}\natural} - G_*$ in Lemma 4 and the second term on the left-hand side of (3.32) in Lemma 5. Thus, we have only to estimate the last term of (3.32). By (2.8), (5.2) and $\sigma \leq \varepsilon_\natural^p$, we have

$$\|B^{Q\natural}\|_* \leq \frac{1}{\nu_{\min}}\|B^{Q\natural}\|_F \leq \frac{1}{\nu_{\min}}\zeta_g^2\sigma^2 \leq \frac{\zeta_g^2\varepsilon_\natural^p}{\nu_{\min}}\sigma. \tag{5.10}$$

Furthermore, by (5.9), it holds that

$$\begin{aligned} \|B^{P\natural} - G_*\|_* &\leq \|B^{P\natural} - B\|_* + \|B - G_*\|_* \\ &\leq K\sigma + \|B - G_*\|_*. \end{aligned} \tag{5.11}$$

By using all the estimates (5.7), (5.8), (5.10) and (5.11), we have

$$\begin{aligned} \|B_+ - G_*\|_* &\leq \|B_+^{\mathrm{DFP}\natural} - G_*\|_* + |\phi - 1|\,\|\Delta B(s, y^{P\natural}, B^{P\natural})\|_* + \|B^{Q\natural}\|_* \\ &\leq (1+\omega_1\sigma)\|B^{P\natural} - G_*\|_* + \omega_2\sigma - \frac{1}{2\delta_\natural}\frac{\|(\widehat{B}^{P\natural} - I)\widehat{s}\|^2}{\|\widehat{s}\|^2} \\ &\qquad + \phi'\tau_1(\delta_\natural)\left[\tau_2(\delta_\natural)\sigma + \frac{\|(\widehat{B}^{P\natural} - I)\widehat{s}\|^2}{\|\widehat{s}\|^2}\right] + \frac{\zeta_g^2\varepsilon_\natural^p}{\nu_{\min}}\sigma \\ &\leq (1+\omega_1\sigma)\|B - G_*\|_* + (1+\omega_1\sigma)K\sigma + \omega_2\sigma + \phi'\tau_1(\delta_\natural)\tau_2(\delta_\natural)\sigma + \frac{\zeta_g^2\varepsilon_\natural^p}{\nu_{\min}}\sigma \\ &\qquad - \frac{1}{2\delta_\natural}\frac{\|(\widehat{B}^{P\natural} - I)\widehat{s}\|^2}{\|\widehat{s}\|^2} + \phi'\tau_1(\delta_\natural)\frac{\|(\widehat{B}^{P\natural} - I)\widehat{s}\|^2}{\|\widehat{s}\|^2} \\ &\leq (1+\mu_1\sigma)\|B - G_*\|_* + \mu_2(\delta_\natural)\sigma - \mu_3(\delta_\natural)\frac{\|(\widehat{B}^{P\natural} - I)\widehat{s}\|^2}{\|\widehat{s}\|^2}, \end{aligned} \tag{5.12}$$

where $\mu_1 = \omega_1 \geq 0$, $\mu_2(\delta_\natural) = (1+\omega_1\varepsilon_\natural^p)K + \omega_2 + \zeta_g^2\varepsilon_\natural^p/\nu_{\min} + \phi'\tau_1(\delta_\natural)\tau_2(\delta_\natural) \geq 0$, $\mu_3(\delta_\natural) = 1/(2\delta_\natural) - \phi'\tau_1(\delta_\natural)$. By the same argument as in the proof of Lemma 7 in Ogasawara and Yabe [10], we can take constants $\mu_2 \geq 0$ and $\mu_3 > 0$ independent of $\delta_\natural$ such that $\mu_2(\delta_\natural) \leq \mu_2$ and $\mu_3(\delta_\natural) \geq \mu_3$. For completeness of the proof of this lemma, we state it here again. It is easy to check that $\mu_3(\delta_\natural) > 0$ if and only if $\delta_\natural < \bar{\delta}$. Since $\gamma \geq 1$ and $\phi' \geq 1$, we must have $\bar{\delta} \leq 1/9$. Clearly, $\tau_1(\delta_\natural)$ and $\tau_2(\delta_\natural)$ are, respectively, increasing and nondecreasing functions in $\delta_\natural < 1$, so $\mu_2(\delta_\natural)$ and $\mu_3(\delta_\natural)$ are, respectively, nondecreasing and decreasing functions in $\delta_\natural \in (0,1)$. Then, by $\delta_\natural \leq 0.9\bar{\delta}$, it follows that $\mu_2(\delta_\natural) \leq \mu_2(0.9\bar{\delta}) = \mu_2$ and $\mu_3(\delta_\natural) \geq \mu_3(0.9\bar{\delta})$. Since $\mu_3(\bar{\delta}) = 0$, we have $1/(2\bar{\delta}) = \phi'\tau_1(\bar{\delta})$. Therefore, $\mu_3(0.9\bar{\delta}) = 1/(1.8\bar{\delta}) - \phi'\tau_1(0.9\bar{\delta}) \geq 1/(1.8\bar{\delta}) - \phi'\tau_1(\bar{\delta}) = 1/(1.8\bar{\delta}) - 1/(2\bar{\delta}) = (1/0.9 - 1)/(2\bar{\delta}) = \mu_3$. Thus, replacing $\mu(\delta_\natural)$ and $\mu_3(\delta_\natural)$ in (5.12) by, respectively, bounds μ_2 and μ_3, we finally obtain the desired result. $\square$

We are ready to prove local and q-linear convergence of the Sheng-Zou-Broyden method. We make a technical assumption to exclude referring to an obvious case. When we consider the sequence of iterates $\{x_k\}$, we always assume for convenience that finite convergence does not occur, i.e., $g_k \neq 0$. This assumption implies that $x_k \neq x_*$ for all $k \geq 0$.

Theorem 1 *Suppose that the standard assumptions* (A1)–(A4) *hold for a nonzero residual problem, i.e.,* $\|r(x_*)\| > 0$. *Assume that there exist constants* $\zeta_\flat \geq 0$ *and* $\varepsilon_\flat > 0$ *such that* $y^\flat$ *satisfies*

$$\|y^\flat - S_* s\| \leq \zeta_\flat \sigma(x, x_+)\|s\| \qquad \text{for all } x, x_+ \in \mathcal{D}_\flat, \tag{5.13}$$

where $s := x_+ - x$, $\mathcal{D}_\flat \equiv \{x \in \mathbf{R}^n : \|x - x_*\| < \varepsilon_\flat\}$. *Let the sequence* $\{x_k\}$ *be generated by the Sheng-Zou-Broyden method with nonnegatively bounded parameters, i.e., there exists a constant* $\phi_U \geq 0$ *such that* $\phi_k \in [0, \phi_U]$ *for all* $k \geq 0$. *Then, for any* $\varrho \in (0,1)$, *there exist positive constants* ε *and* δ *such that if*

$$\|x_0 - x_*\| < \varepsilon, \qquad \|(J_0 + W_0)^\top (J_0 + W_0) - G_*\|_* < \delta, \qquad W_0^\top r_0 = 0,$$

the sequence $\{x_k\}$ *is well-defined and converges at least q-linearly to* x_* *with a rate of convergence* ϱ, *i.e.,*

$$\|x_{k+1} - x_*\| \leq \varrho \|x_k - x_*\|, \quad k = 0, 1, \ldots.$$

Proof. The proof of this theorem is along the same line as the proof of Theorem 1 in Ogasawara and Yabe [10]. By assumptions, Lemma 3 ensures that there exist constants $\widehat{\zeta}_\natural \geq 0$, $\gamma \geq 1$ and $\varepsilon_\natural > 0$ such that (5.5) and (5.6) hold for all distinct $x, x_+ \in \mathcal{D}_\natural = \{x \in \mathbf{R}^n : \|x - x_*\| < \varepsilon_\natural\} \subset \mathcal{D}_* \cap \mathcal{D}_c \cap \mathcal{D}_\flat$. Define the set

$$\mathcal{N} \equiv \{B \in \mathbf{R}^{n \times n} : \|B - G_*\|_* < 2\delta\}.$$

Let $\varrho \in (0,1)$ be fixed. We prove by induction on k that $(\mathrm{E1})_k$–$(\mathrm{E4})_k$ hold for all $k \geq 0$:

$(\mathrm{E1})_k$ $$B_k \in \mathcal{N}, \qquad \|B_k\| \leq K_1, \qquad \|B_k^{-1}\| \leq K_2, \qquad \|W_k\| \leq K_3,$$

$(\mathrm{E2})_k$ $$\|x_{k+1} - x_*\| \leq \varrho \|x_k - x_*\|, \quad x_{k+1} \in \mathcal{D}_\natural, \quad x_{k+1} \neq x_k,$$

$(\mathrm{E3})_k$ $$\|B_k^{P\natural} - B_k\|_* \leq K_4 \sigma_k,$$

$(\mathrm{E4})_k$ $$\|B_{k+1} - G_*\|_* \leq (1 + \mu_1 \sigma_k)\|B_k - G_*\|_* + \mu_2 \sigma_k - \mu_3 \frac{\|(\widehat{B}_k^{P\natural} - I)\widehat{s}_k\|^2}{\|\widehat{s}_k\|^2},$$

where positive constants K_1, K_2, K_3 and K_4 are defined by

$$K_1 := 1.1\nu_{\max}, \quad K_2 := \frac{1}{\nu_{\min}} + \frac{\varrho}{\nu_{\max}},$$

$$K_3 := \|J_*\| + \sqrt{K_1} + \xi_J \varepsilon_\natural^p, \quad K_4 := \frac{2\sqrt{n}}{\nu_{\min}}(\sqrt{K_1} + \xi_J \varepsilon_\natural^p)(2\xi_J + \zeta_g + \zeta_r K_3),$$

$\mathcal{D}_\natural$, $\varepsilon_\natural$ are given in Lemma 3, $\mu_1 \geq 0$, $\mu_3 > 0$ are given in Lemma 6, and $\mu_2 \geq 0$ is also given in Lemma 6 but with K_4 replacing K.

Choose $\delta = \delta(\varrho)$ and $\varepsilon = \varepsilon(\varrho)$ such that

$$\varepsilon \leq \varepsilon_\natural, \tag{5.14}$$

$$\frac{1}{\nu_{\max}} \cdot \frac{\xi_G}{p+1}\varepsilon^p + 2\delta \leq \frac{\varrho}{\kappa_* + \varrho}, \tag{5.15}$$

$$K_4 \varepsilon^p \leq 0.1\delta, \tag{5.16}$$

$$\delta \leq \frac{3}{7}\bar{\delta}, \tag{5.17}$$

$$\frac{(2\delta\mu_1 + \mu_2)}{1 - \varrho^p}\varepsilon^p \leq \delta, \tag{5.18}$$

where $\kappa_* \equiv \nu_{\max}/\nu_{\min} \geq 1$, and $\bar{\delta}$ is given in Lemma 6. Recall that $\bar{\delta} \leq 1/9$.

We first consider the case $k = 0$.

(E1)$_0$ It is clear from the choice of the initial matrix that $B_0 \in \mathcal{N}$. Hence, it follows from (2.6), (2.8) and (5.17) that

$$\begin{aligned} \|B_0\| &\leq \|B_0 - G_*\| + \|G_*\| \leq \nu_{\max}\|B_0 - G_*\|_* + \nu_{\max} \\ &< (2\delta + 1)\nu_{\max} \leq \Big(\frac{6}{7}\bar{\delta} + 1\Big)\nu_{\max} \\ &\leq \Big(\frac{6}{7}\cdot\frac{1}{9} + 1\Big)\nu_{\max} = \Big(\frac{2}{21} + 1\Big)\nu_{\max} < \Big(\frac{2}{20} + 1\Big)\nu_{\max} = K_1. \end{aligned} \tag{5.19}$$

We recall here a variant of the so-called Banach perturbation lemma (see Theorem 3.1.4 in Dennis and Schnabel [7]): If A is symmetric positive definite and $\|A^{-1/2}(B-A)A^{-1/2}\| < 1$, then B is nonsingular and

$$\|B^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1/2}(B-A)A^{-1/2}\|}. \tag{5.20}$$

Note that from (5.15), we get

$$2\delta \leq \frac{\varrho}{\kappa_* + \varrho}.$$

Then, we have from $B_0 \in \mathcal{N}$ and the above inequality that

$$\begin{aligned} \|G_*^{-1/2}(B_0 - G_*)G_*^{-1/2}\| &\leq \|G_*^{-1/2}(B_0 - G_*)G_*^{-1/2}\|_F \\ &= \|B_0 - G_*\|_* \\ &< 2\delta \leq \frac{\varrho}{\kappa_* + \varrho} < \frac{1}{2}. \end{aligned} \tag{5.21}$$

Therefore, by the Banach perturbation lemma, B_0 is also nonsingular and, from (5.20) and (5.21) together with (2.6), we have

$$\begin{aligned} \|B_0^{-1}\| &\leq \frac{\|G_*^{-1}\|}{1 - \|G_*^{-1/2}(B_0 - G_*)G_*^{-1/2}\|} \\ &\leq \frac{1}{\nu_{\min}} \cdot \frac{1}{1 - \frac{\varrho}{\kappa_* + \varrho}} = \frac{1}{\nu_{\min}} \cdot \frac{\kappa_* + \varrho}{\kappa_*} \\ &= \frac{\kappa_* + \varrho}{\nu_{\max}} = \frac{1}{\nu_{\min}} + \frac{\varrho}{\nu_{\max}} = K_2. \end{aligned} \tag{5.22}$$

By using (2.2) of assumption (A3), (5.14) and (5.19), we have

$$\begin{aligned} \|W_0\| &\leq \|J_0 + W_0\| + \|J_* - J_0\| + \|-J_*\| \\ &\leq \|B_0\|^{1/2} + \xi_J\|x_0 - x_*\|^p + \|J_*\| \\ &\leq \sqrt{K_1} + \xi_J\varepsilon_\natural^p + \|J_*\| = K_3. \end{aligned} \tag{5.23}$$

(E2)$_0$ Obviously, from the choice of the initial point and (5.14), we have $x_0 \in \mathcal{D}_\natural$. By (2.5), (2.8) and (5.22), we obtain

$$\begin{aligned} \|x_1 - x_*\| &= \|(x_0 - B_0^{-1}g_0) - x_*\| \\ &= \Big\|-B_0^{-1}\{g_0 - g_* - (G_* - G_* + B_0)(x_0 - x_*)\}\Big\| \\ &\leq \|B_0^{-1}\|(\|g_0 - g_* - G_*(x_0 - x_*)\| + \|B_0 - G_*\|\|x_0 - x_*\|) \\ &< \frac{\kappa_* + \varrho}{\nu_{\max}}\Big(\frac{\xi_G}{p+1}\varepsilon^p + 2\nu_{\max}\delta\Big)\|x_0 - x_*\| \\ &\leq \varrho\|x_0 - x_*\| \leq \varrho\varepsilon < \varepsilon_\natural, \end{aligned}$$

where the third inequality comes from (5.15). This implies that $x_1 \in \mathcal{D}_\natural$ and $x_1 \neq x_0$, because $x_k \neq x_*$ for all $k \geq 0$.

$(\mathrm{E3})_0$ Recalling (3.20), and then using (4.4), (2.2), (5.2) and (5.3) together with (5.14), (5.19) and (5.23), we see that

$$\begin{aligned}
\|B_0^{P\natural} - B_0\| &= \|(P_0J_1 + P_0W_0)^\top(P_0J_1 + P_0W_0) - (J_0 + W_0)^\top(J_0 + W_0)\| \\
&\leq \|P_0J_1 + J_0 + (P_0 + I)W_0\|\|P_0J_1 - J_0 + (P_0 - I)W_0\| \\
&= \|P_0(J_1 - J_0) + (P_0 + I)(J_0 + W_0)\|\|J_1 - J_0 - Q_0(J_1 + W_0)\| \\
&\leq (\|J_1 - J_*\| + \|J_* - J_0\| + 2\|J_0 + W_0\|) \\
&\quad \times(\|J_1 - J_*\| + \|J_* - J_0\| + \|Q_0J_1\| + \|Q_0W_0\|) \\
&\leq \{\xi_J(\|x_1 - x_*\|^p + \|x_0 - x_*\|^p) + 2\|B_0\|^{1/2}\} \\
&\quad \times\{\xi_J(\|x_1 - x_*\|^p + \|x_0 - x_*\|^p) + \zeta_g\sigma_0 + \zeta_r\|W_0\|\sigma_0)\} \\
&\leq (2\xi_J\varepsilon_\natural^p + 2\sqrt{K_1}) \cdot (2\xi_J + \zeta_g + \zeta_r K_3)\sigma_0 = \frac{\nu_{\min}}{\sqrt{n}} K_4\sigma_0. \qquad (5.24)
\end{aligned}$$

Similarly, from (4.5) and (5.24), we have

$$\begin{aligned}
\|B_0^{P\natural} - B_0\|_F &\leq \sqrt{n}\,\|P_0J_1 + J_0 + (P_0 + I)W_0\|\|P_0J_1 - J_0 + (P_0 - I)W_0\| \\
&\leq \sqrt{n} \cdot \frac{\nu_{\min}}{\sqrt{n}} K_4\sigma_0 = \nu_{\min}K_4\sigma_0.
\end{aligned}$$

Hence, it follows from (2.8) that

$$\|B_0^{P\natural} - B_0\|_* \leq \frac{1}{\nu_{\min}}\|B_0^{P\natural} - B_0\|_F \leq K_4\sigma_0.$$

$(\mathrm{E4})_0$ Since by (5.16)

$$\|B_0^{P\natural} - B_0\|_* \leq K_4\sigma_0 \leq K_4\varepsilon^p \leq 0.1\delta,$$

we have with (5.17)

$$\begin{aligned}
\|B_0^{P\natural} - G_*\|_* &\leq \|B_0^{P\natural} - B_0\|_* + \|B_0 - G_*\|_* \\
&< 0.1\delta + 2\delta = 2.1\delta \leq 2.1 \cdot \frac{3}{7}\bar{\delta} = 0.9\bar{\delta}.
\end{aligned}$$

Thus, since W_1 is well-defined, we can apply Lemma 6 to $K = K_4$ and $\delta_\natural = 2.1\delta$ to obtain

$$\|B_1 - G_*\|_* \leq (1 + \mu_1\sigma_0)\|B_0 - G_*\|_* + \mu_2\sigma_0 - \mu_3\frac{\|(\widehat{B}_0^{P\natural} - I)\widehat{s}_0\|^2}{\|\widehat{s}_0\|^2}.$$

Therefore, we have shown the case $k = 0$.

Now assume that expressions $(\mathrm{E1})_k$–$(\mathrm{E4})_k$ hold for $k = 0, \ldots, j$. We want to prove that they are also valid for $k = j + 1$. We will show below only the first statement of $(\mathrm{E1})_{j+1}$, i.e., $B_{j+1} \in \mathcal{N}$, because the remaining results of $(\mathrm{E1})_{j+1}$–$(\mathrm{E4})_{j+1}$ can be proved by using identical arguments for the case $k = 0$. It follows from $(\mathrm{E4})_k$ and $B_k \in \mathcal{N}$ of $(\mathrm{E1})_k$ that

$$\|B_{k+1} - G_*\|_* + \mu_3\frac{\|(\widehat{B}_k^{P\natural} - I)\widehat{s}_k\|^2}{\|\widehat{s}_k\|^2} \leq \|B_k - G_*\|_* + (2\delta\mu_1 + \mu_2)\sigma_k \qquad (5.25)$$

for all $k = 0, \ldots, j$. Since

$$\sigma_k = \|x_k - x_*\|^p \le (\varrho^k \|x_0 - x_*\|)^p \le (\varrho^p)^k \varepsilon^p,$$

by summing both sides of (5.25) from $k = 0$ to j, and using (5.18), we have

$$\begin{aligned} \|B_{j+1} - G_*\|_* + \mu_3 \sum_{k=0}^{j} \frac{\|(\widehat{B}_k^{P\natural} - I)\widehat{s}_k\|^2}{\|\widehat{s}_k\|^2} &\le \|B_0 - G_*\|_* + (2\delta\mu_1 + \mu_2) \sum_{k=0}^{j} \sigma_k \\ &< \delta + (2\delta\mu_1 + \mu_2) \frac{\varepsilon^p}{1 - \varrho^p} \\ &\le 2\delta. \end{aligned} \tag{5.26}$$

This implies that $\|B_{j+1} - G_*\|_* < 2\delta$, i.e., $B_{j+1} \in \mathcal{N}$. □

We can now establish the main result of this paper, which states q-superlinear convergence of the Sheng-Zou-Broyden method.

Theorem 2 *Suppose that all conditions of Theorem* 1 *hold. Then, the sequence* $\{x_k\}$ *generated by the Sheng-Zou-Broyden method converges q-superlinearly to* x_*.

Proof. The proof of this theorem is essentially the same as the proof of Theorem 2 of Ogasawara and Yabe [10]. Since we also have from (5.26) that for any $j \ge 0$

$$\mu_3 \sum_{k=0}^{j} \frac{\|(\widehat{B}_k^{P\natural} - I)\widehat{s}_k\|^2}{\|\widehat{s}_k\|^2} < 2\delta,$$

we find that for $j \to \infty$ the series on the left-hand side converges, and, in particular, that

$$\lim_{k\to\infty} \frac{\|(\widehat{B}_k^{P\natural} - I)\widehat{s}_k\|}{\|\widehat{s}_k\|} = 0. \tag{5.27}$$

Similar to (5.24), we can deduce that for all $k \ge 0$

$$\|B_k^{P\natural} - B_k\| \le \frac{\nu_{\min}}{\sqrt{n}} K_4 \sigma_k. \tag{5.28}$$

Using (5.28) and (2.6), we have

$$\begin{aligned} \frac{\|(B_k - G_*)s_k\|}{\|s_k\|} &\le \frac{\|(B_k - B_k^{P\natural})s_k\|}{\|s_k\|} + \frac{\|(B_k^{P\natural} - G_*)s_k\|}{\|s_k\|} \\ &\le \|B_k - B_k^{P\natural}\| + \frac{\|G_*^{1/2} s_k\|}{\|s_k\|} \frac{\|G_*^{1/2}(\widehat{B}_k^{P\natural} - I)G_*^{1/2} s_k\|}{\|G_*^{1/2} s_k\|} \\ &\le \frac{\nu_{\min}}{\sqrt{n}} K_4 \sigma_k + \nu_{\max} \frac{\|(\widehat{B}_k^{P\natural} - I)\widehat{s}_k\|}{\|\widehat{s}_k\|}. \end{aligned}$$

Therefore, by $\lim_{k\to\infty} \sigma_k = 0$ and (5.27), we obtain

$$\lim_{k\to\infty} \frac{\|(B_k - G_*)s_k\|}{\|s_k\|} = 0.$$

This is the well-known Dennis-Moré characterization [6] for a convergent sequence $\{x_k\}$ to converge q-superlinearly to x_*. □

Remark 2 It is known that the Dennis-Biggs vector $y^\flat = y^{\mathrm{DB}\flat}$ given by (3.13) satisfies (5.13). See Lemma 4.1 of Dennis, Martínez and Tapia [5]. See also Lemma 8 of Ogasawara and Yabe [10].

6. Concluding Remarks

In this paper, we have shown local and q-superlinear convergence of the Sheng-Zou-Broyden method with W-updates from the nonnegatively bounded class for *nonzero* residual problems. This result is a generalization of that for the Sheng-Zou-BFGS method stated without a full proof by Sheng and Zou [12]. Our convergence analysis heavily depends on the assumption of nonzero residuals, so the arguments presented in this paper are not able to be carried over to the zero residual case. At present, we have no idea how we can prove or disprove superlinear or even local convergence of the method for zero residual problems.

The assumption of nonzero residuals is, however, not so restrictive at least in practice. One reason is that most (but not all) of nonlinear least squares problems may have nonzero residuals. Another reason is that we can apply any of two strategies proposed for zero residual problems, i.e., sizing and hybridizing techniques. The sizing techniques are multiplying the approximating matrix by a scalar before updating. Sizing factors for A-updates were proposed by Bartholomew-Biggs [2], and Dennis, Gay and Welsch [4]. Sizing factors for W-updates were proposed by Yabe and Takahashi [17], [18]. The sizing strategies have been found to work effectively, and the factorized/structured quasi-Newton methods incorporated with sizing perform well in practice on zero residual problems as well. For details about numerical results, see the references cited above. On the other hand, the hybridizing techniques are the combination of the Gauss-Newton method and the (un)structured quasi-Newton method, and the switching between these two methods. The hybrid methods were proposed by Al-Baali and Fletcher [1] and Fletcher and Xu [8]. See these references.

Although the sizing and hybridizing strategies are practical means as a remedy, the issue of analyzing the convergence behavior of the Sheng-Zou-Broyden method on *zero* residual problems is also theoretically an interesting and challenging problem.

Acknowledgments

The authors would like to thank anonymous referees for their valuable comments to improve the paper.

References

[1] M. Al-Baali and R. Fletcher: Variational methods for non-linear least squares. *Journal of the Operational Research Society*, **36** (1985) 405–421.

[2] M. C. Bartholomew-Biggs: The estimation of the Hessian matrix in nonlinear least squares problems with non-zero residuals. *Mathematical Programming*, **12** (1977) 67–80.

[3] J. E. Dennis, Jr.: A brief survey of convergence results for quasi-Newton methods. In R. Cottle and C. Lemke (eds.): *Nonlinear Programming: SIAM-AMS Proceedings, Vol. 9* (American Mathematical Society, Providence, 1976), 185–199.

[4] J. E. Dennis, Jr., D. M. Gay and R. E. Welsch: An adaptive nonlinear least-squares algorithm. *ACM Transactions on Mathematical Software*, **7** (1981) 348–368.

[5] J. E. Dennis, Jr., H. J. Martínez and R. A. Tapia: Convergence theory for the structured BFGS secant method with an application to nonlinear least squares. *Journal of Optimization Theory and Applications*, **61** (1989) 161–178.

[6] J. E. Dennis, Jr. and J. J. Moré: A characterization of superlinear convergence and its application to quasi-Newton methods. *Mathematics of Computation*, **28** (1974) 549–560.

[7] J. E. Dennis, Jr. and R. B. Schnabel: *Numerical Methods for Unconstrained Optimization and Nonlinear Equations* (Prentice-Hall, New Jersey, 1983).

[8] R. Fletcher and C. Xu: Hybrid methods for nonlinear least squares. *IMA Journal of Numerical Analysis*, **7** (1987) 371–389.

[9] H. Ogasawara: A note on the equivalence of a class of factorized Broyden families for nonlinear least squares problems. *Linear Algebra and Its Applications*, **297** (1999) 183–191.

[10] H. Ogasawara and H. Yabe: Superlinear convergence of factorized structured quasi-Newton methods for nonlinear optimization. *Asia-Pacific Journal of Operational Research*, **17** (2000) 55–80.

[11] S. S. Oren and D. G. Luenberger: Self-scaling variable metric (SSVM) algorithms, part I. Criteria and sufficient conditions for scaling a class of algorithms. *Management Science*, **20** (1974) 845–862.

[12] S. -B. Sheng and Z. -H. Zou: A new secant method for nonlinear least squares problems. Technical Report NANOG-1988-03, March 1988, Department of Mathematics, Nanjing University, Nanjing, People's Republic of China. (See also: *Numerical Mathematics, A Journal of Chinese Universities*, **2** (1993) 125–137.)

[13] C. X. Xu, X. F. Ma and M. Y. Kong: A class of factorized quasi-Newton methods for nonlinear least squares problems. *Journal of Computational Mathematics*, **14** (1996) 143–158.

[14] H. Yabe: A family of variable-metric methods with factorized expressions. *TRU Mathematics*, **17** (1981) 141–152.

[15] H. Yabe: Variations of structured Broyden families for nonlinear least squares problems. *Optimization Methods and Software*, **2** (1993) 107–144.

[16] H. Yabe and T. Takahashi: Structured quasi-Newton methods for nonlinear least squares problems. *TRU Mathematics*, **24** (1988) 195–209.

[17] H. Yabe and T. Takahashi: Factorized quasi-Newton methods for nonlinear least squares problems. *Mathematical Programming*, **51** (1991) 75–100.

[18] H. Yabe and T. Takahashi: Numerical comparison among structured quasi-Newton methods for nonlinear least squares problems. *Journal of the Operations Research Society of Japan*, **34** (1991) 287–305.

[19] H. Yabe and N. Yamaki: Convergence of a factorized Broyden-like family for nonlinear least squares problems. *SIAM Journal on Optimization*, **5** (1995) 770–791.

[20] N. Yamaki and H. Yabe: On factorized variable-metric methods. *TRU Mathematics*, **17** (1981) 285–294.

Hideho Ogasawara
Department of Mathematical Information Science
Faculty of Science
Tokyo University of Science
1-3 Kagurazaka, Shinjuku-ku
Tokyo 162-0825, Japan
E-mail: `hoga@rs.kagu.tus.ac.jp`