数理解析研究所講究録
第676巻 1988年 170-194

FACTORIZED QUASI-NEWTON METHODS FOR NONLINEAR LEAST SQUARES PROBLEMS
(非線形最小2乗問題に対する分解型準ニュートン法)

鹿島建設(株)・情報システム部　高橋俊彦 (Toshihiko Takahashi)
東京理科大学　・　工学部　矢部 博 (Hiroshi Yabe)

1. Introduction

We consider the following nonlinear least squares problem:

(1.1) Minimize $F(x) = (1/2) \sum_{j=1}^{m} (f_j(x))^2$, $x \in R^n$, $m \geqq n$,

where each $f_j: R^n \to R$ is twice continuously differentiable. This problem is extremely important in many fields of mathematical programming applications, e.g. maximum likelifood estimations, nonlinear data fitting or parameter estimation, respectively.

Most iterative methods for the above problem are variants of Newton's method. At the k-th iteration of Newton's method, the search direction d_k is computed by

(1.2) $\nabla^2 F(x_k)\, d_k = -\nabla F(x_k)$,

and the new point is generated by

(1.3) $x_{k+1} = x_k + d_k$.

Here x_k is the current estimate of the minimum point x^*, and $\nabla F, \nabla^2 F$ are the gradient vector and the Hessian matrix of F, respectively, and are given by

(1.4) $\nabla F(x) = J(x)^T f(x)$,

(1.5) $\nabla^2 F(x) = J(x)^T J(x) + \sum_{j=1}^{m} f_j(x) \nabla^2 f_j(x)$,

where

(1.6) $f(x) = (f_1(x), \ldots, f_m(x))^T$,

and J is the m x n Jacobian matrix of f, and the symbol "T" denotes the transpose of a vector or a matrix.

Since the cost of providing the complete Hessian matrix is often expensive, some methods have been derived which use only the first derivative information. For example, the Gauss-Newton method and the Levenberg-Marquardt method are well known. These methods neglect the second part of the Hessian matrix of F, so they can be expected to perform well when the residuals at x^* are small or each f_j is close to linear. However,

they can be much less efficient when the neglected part of $\nabla^2 F(x)$ is not small compared with $J(x)^T J(x)$ in the sense of Meyer's result [9].

On the other hand, quasi-Newton approximations to the second part of the Hessian matrix have been considered [7]. Recently, two robust algorithms have been proposed by Bartholomew-Biggs[1] and Dennis, Gay and Welsch [8]. These methods are shown in Section 2. Our approach is based on the idea of structured quasi-Newton updating which utilizes the structure of the Hessian matrix of F. The main purpose of this paper is to obtain descent search directions for the objective function, which may enable us to establish global convergence property under suitable conditions. Subsequently, to accomplish the above desirable property, we propose factorized versions of the structured quasi-Newton methods and derive various types of factorized quasi-Newton updating formulae in Section 3. Consequently, in Section 4, we prove the local and q-superlinear convergence of our algorithms. Finally, some computational experiments are given in order to show that our methods are comparable to other effective methods.

Throughout this paper, the norm $\|\cdot\|$ denotes the 2-norm for vectors and matrices. For any matrix Q and a nonsingular M, $\|Q\|_F$ and $\|Q\|_{F,M}$ denote the Frobenius and the weighted Frobenius norms of Q, respectively, and are defined by

$$\|Q\|_F = (\mathrm{Trace}(QQ^T))^{1/2} \quad \text{and} \quad \|Q\|_{F,M} = \|M^T Q M\|_F. \tag{1.7}$$

For a symmetric positive definite matrix Q, $Q^{1/2}$ denotes the symmetric matrix which satisfies $(Q^{1/2})^2 = Q$.

2. Structured Quasi-Newton Methods for Nonlinear Least Squares Problems

A straightforward application of general quasi-Newton methods to the nonlinear least squares problem is not desirable, because these methods approximate all of the Hessian matrix. Since the nonlinear least squares algorithms usually calculate the Jacobian matrix J(x) analytically or numerically, the portion $J(x)^T J(x)$ of $\nabla^2 F(x)$ is always readily available, so we only have to approximate the second part of $\nabla^2 F(x)$. Therefore, for the nonlinear least squares problem, it has been considered that the search direction can be computed by

$$(J_k^T J_k + A_k) d_k = -J_k^T f_k, \tag{2.1}$$

where $f_k = f(x_k)$, $J_k = J(x_k)$, and the matrix A_k is the k-th approximation to

the second part of the Hessian matrix of F [7]. The matrix A_k is updated such that the new matrix A_{k+1} satisfies the secant condition

$$(2.2) \qquad A_{k+1} s_k = u_k, \quad u_k = y_k - J_{k+1}^T J_{k+1} s_k$$

or

$$(2.3) \qquad A_{k+1} s_k = v_k, \quad v_k = (J_{k+1} - J_k)^T f_{k+1},$$

where

$$(2.4) \qquad s_k = x_{k+1} - x_k, \quad y_k = \nabla F_{k+1} - \nabla F_k, \quad \nabla F_k = \nabla F(x_k).$$

The first is proposed by Broyden and Dennis (BD) [4], and the second is proposed by Bartholomew-Biggs(Biggs)[1] and Dennis, Gay and Welsch (DGW)[8]. We call these strategies structured quasi-Newton methods.

Structured quasi-Newton updates are usually of rank one or of rank two. Broyden and Dennis gave the following update:

(i)the BD update

$$(2.5) \qquad A_{k+1} = A_k + ((u_k - A_k s_k) s_k^T + s_k (u_k - A_k s_k)^T)/s_k^T s_k - ((u_k - A_k s_k)^T s_k / (s_k^T s_k)^2) s_k s_k^T.$$

Recently, by using sizing techniques, Bartholomew-Biggs and Dennis et al. have proposed the robust algorithms for the both cases of large and small residual problems. When the residuals are large, their algorithms perform as well as the Broyden and Dennis method does. On the other hand, for the very small residual problems, their algorithms perform almost as well as the Gauss-Newton method does. Their updates are as follows:

(ii)the Biggs update

$$(2.6) \qquad A_{k+1} = \beta_k A_k + (v_k - \beta_k A_k s_k)(v_k - \beta_k A_k s_k)^T / (v_k - \beta_k A_k s_k)^T s_k,$$

$$(2.7) \qquad \beta_k = | f_{k+1}^T f_k / f_k^T f_k |,$$

(iii)the DGW update

$$(2.8) \qquad A_{k+1} = \beta_k A_k + ((v_k - \beta_k A_k s_k) y_k^T + y_k (v_k - \beta_k A_k s_k)^T)/s_k^T y_k - (s_k^T (v_k - \beta_k A_k s_k)/(s_k^T y_k)^2) y_k y_k^T,$$

$$(2.9) \qquad \beta_k = \min(| s_k^T v_k / s_k^T A_k s_k | , 1),$$

where β_k is a sizing factor.

3. Factorized Versions of Structured Quasi-Newton Methods

In order to obtain a descent search direction, it is desirable that the

coefficient matrix in (2.1) is positive definite. However, it is not clear how to construct updating formulae of A_k such that the matrix $J_k^T J_k + A_k$ is positive definite. To overcome this difficulty, several strategies have been proposed,for example, the modified Cholesky decomposition of the matrix $J_k^T J_k + A_k$, the Levenberg-Marquardt modification (the model/trust region strategy)[8] or switching to the Gauss-Newton method [2].

In this section, a direct approach is proposed in order to maintain positive definiteness of the coefficient matrix in (2.1). We try to compute the search direction by solving the linear system of equations

$$(3.1) \qquad (L_k + J_k)^T (L_k + J_k)\, d_k = -J_k^T f_k,$$

where the matrix L_k is an $m \times n$ correction matrix to the Jacobian matrix such that $L_k^T L_k + L_k^T J_k + J_k^T L_k$ is the k-th approximation to the second part of the Hessian matrix of F [10]. Since the coefficient matrix is expressed by the factorized form, the search direction may be expected to be a descent direction for F.

Now we construct updating formulae of the matrix L_k. The secant condition (2.2) or (2.3) for A_{k+1} can be reduced to the following secant condition for L_{k+1}

$$(3.2) \qquad (L_{k+1} + J_{k+1})^T (L_{k+1} + J_{k+1})\, s_k = z_k,$$

where

$$(3.3) \qquad z_k = y_k$$

or

$$(3.4) \qquad z_k = v_k + J_{k+1}^T J_{k+1} s_k,$$

and the vectors v_k, s_k and y_k are given in (2.3) and (2.4), respectively. It is easily shown that, for nonzero s_k and z_k, the matrix equation (3.2) is consistent if and only if

$$(3.5) \qquad L_{k+1}^T h = z_k - J_{k+1}^T h \quad \text{and} \quad L_{k+1} s_k = h - J_{k+1} s_k$$

for some m-dimensional vector h. Further, the matrix equations (3.5) have a common solution L_{k+1} if and only if each equation separately has a solution and $h^T h = s_k^T z_k$.

In the sequent two subsections, we find a rectangular matrix L_{k+1} which satisfies the equations (3.5) under the assumption of $s_k^T z_k > 0$.

3.1 Least-change secant updates of L_k

Since the equations (3.5) may not uniquely determine the solution matrix L_{k+1}, we use the least-change secant update technique following to Dennis and Schnabel [6].

For any unknown m-dimensional vector h such that $h^T h = s_k^T z_k$, minimizing the Frobenius norm

(3.6) $\quad \| L_{k+1}^T - L_k^T \|_F$

with respect to L_{k+1}, subject to

(3.7) $\quad L_{k+1}^T h = z_k - J_{k+1}^T h,$

we have

$$L_{k+1} = L_k + h(z_k - (L_k + J_{k+1})^T h)^T / h^T h.$$

By substituting the above for the other condition in (3.5) and using $h^T h = s_k^T z_k$, the vector h can be determined by the form

$$h = (s_k^T z_k / s_k^T B_k^{\#} s_k)^{1/2} (L_k + J_{k+1}) s_k,$$

where

(3.8) $\quad B_k^{\#} = (L_k + J_{k+1})^T (L_k + J_{k+1}).$

Thus we have the rank one update of L_k as follows:

(3.9) $\quad L_{k+1} = L_k + ((L_k + J_{k+1}) s_k / s_k^T B_k^{\#} s_k)((s_k^T B_k^{\#} s_k / s_k^T z_k)^{1/2} z_k - B_k^{\#} s_k)^T.$

Setting

(3.10) $\quad B_{k+1} = (L_{k+1} + J_{k+1})^T (L_{k+1} + J_{k+1}),$

we have

(3.11) $\quad B_{k+1} = B_k^{\#} - B_k^{\#} s_k s_k^T B_k^{\#} / s_k^T B_k^{\#} s_k + z_k z_k^T / s_k^T z_k,$

which is the analogy of the BFGS update. Note that (3.11) differs from the standard BFGS update in that the structure of the Hessian matrix (1.5) is included in z_k and $B_k^{\#}$.

Next, we consider an analogy of the DFP update. For any unknown m-dimensional vector h, which satisfies $h^T h = s_k^T z_k$, and the nonsingular matrices $W_L \in R^{m \times m}$, $W_R \in R^{n \times n}$, minimizing the Frobenius norm

(3.12) $\quad \| W_L (L_{k+1} - L_k) W_R \|_F$

with respect to L_{k+1}, subject to

(3.13) $L_{k+1}s_k = h - J_{k+1}s_k$,

we have

$$L_{k+1} = L_k + (h - (L_k + J_{k+1})s_k)s_k^T W / s_k^T W s_k,$$

where $W = W_R^{-T}W_R^{-1}$. If the symmetric positive definite matrix W is chosen such that $Ws_k = z_k$, by substituting the above for the other condition in (3.5), we have

$$(L_k + J_{k+1})^T h = c\, z_k, \qquad c = s_k^T(L_k + J_{k+1})^T h / s_k^T z_k.$$

If the matrix $L_k + J_{k+1}$ is of full rank, then

$$h = c\,((L_k + J_{k+1})^T)^+ z_k,$$

where $((L_k+J_{k+1})^T)^+$ is the Moore-Penrose generalized inverse of $(L_k+J_{k+1})^T$. Setting $h^T h = s_k^T z_k$, we have the rank one update of L_k as follows:

(3.14) $$L_{k+1} = L_k + (L_k+J_{k+1})((s_k^T z_k / z_k^T (B_k^{\#})^{-1} z_k)^{1/2}(B_k^{\#})^{-1} z_k - s_k)(z_k / s_k^T z_k)^T.$$

Further, we have

(3.15) $$B_{k+1} = B_k^{\#} + (1 + s_k^T B_k^{\#} s_k / s_k^T z_k) z_k z_k^T / s_k^T z_k - (B_k^{\#} s_k z_k^T + z_k s_k^T B_k^{\#}) / s_k^T z_k,$$

which is the analogy of the DFP update.

3.2 Least-change secant updates of $L_k + J_k$

In order to obtain the updating formulae which satisfy the secant condition (3.2), we minimized the norm (3.6) or (3.12) in the previous subsection. Instead of this strategy, we can minimize the norms

(3.16) $$\| (L_{k+1} + J_{k+1})^T - (L_k + J_k)^T \|_F,$$

and

(3.17) $$\| W_L((L_{k+1} + J_{k+1}) - (L_k + J_k))W_R \|_F.$$

Then, by the same way as the subsection 3.1, we have the following updates:
(i) the update corresponding to (3.9)

(3.18) $$L_{k+1} = L_k + J_k - J_{k+1} + ((L_k+J_k)s_k / s_k^T B_k s_k)((s_k^T B_k s_k / s_k^T z_k)^{1/2} z_k - B_k s_k)^T,$$

(3.19) $$B_{k+1} = B_k - B_k s_k s_k^T B_k / s_k^T B_k s_k + z_k z_k^T / s_k^T z_k,$$

(ii)the update corresponding to (3.14)

(3.20) $$L_{k+1} = L_k + J_k - J_{k+1} + (L_k+J_k)((s_k^T z_k / z_k^T B_k^{-1} z_k)^{1/2} B_k^{-1} z_k - s_k)(z_k / s_k^T z_k)^T,$$

(3.21) $$B_{k+1} = B_k + (1 + s_k^T B_k s_k / s_k^T z_k) z_k z_k^T / s_k^T z_k - (B_k s_k z_k^T + z_k s_k^T B_k) / s_k^T z_k,$$

where $B_k = (L_k + J_k)^T(L_k + J_k)$. For the case of $z_k = y_k$, we have the standard BFGS and DFP updates, which are identical with the results of Dennis and Schnabel [6]. For the case of $z_k = v_k + J_{k+1}^T J_{k+1} s_k$, we have another updating formulae.

For (3.19) and (3.21), the inverse updating formulae can be obtained by letting $H_k = B_k^{-1}$. Then we have the inverse updates as follows:

(i)' the inverse update corresponding to (3.19)

$$(3.22) \quad H_{k+1} = H_k + (1+z_k^T H_k z_k / s_k^T z_k) s_k s_k^T / s_k^T z_k - (H_k z_k s_k^T + s_k z_k^T H_k)/s_k^T z_k,$$

(ii)' the inverse update corresponding to (3.21)

$$(3.23) \quad H_{k+1} = H_k - H_k z_k z_k^T H_k / z_k^T H_k z_k + s_k s_k^T / s_k^T z_k.$$

In this case, the search direction can be calculated by

$$(3.24) \quad d_k = -H_k J_k^T f_k,$$

without solving the linear system of equations.

3.3. Sizing of the updating matrix

We know that, for zero residual problems, the matrices A_k and $L_k^T L_k + L_k^T J_k + J_k^T L_k$ should ideally converge to zero. If the matrices do not at least become small in those cases, then structured quasi-Newton methods cannot be hoped to compete with the Gauss-Newton method. Noting that the quasi-Newton updates do not generate the zero matrix, some remedies must be employed. Among them, the sizing of the updating matrices which has been introduced by Bartholomew-Biggs[1] and Dennis et al.[8] seems most promising. The Biggs' sizing factor (2.7) is based on the idea such that if $f_{k+1} = \beta_k f_k$ for some β_k, $A_k = \sum_{i=1}^m f_i^k \nabla^2 f_i^k$ and each f_i is quadratic, then $\sum_{i=1}^m f_i^{k+1} \nabla^2 f_i^{k+1} = \beta_k A_k$, where f_i^{k+1} and $\nabla^2 f_i^{k+1}$ denote $f_i(x_{k+1})$ and $\nabla^2 f_i(x_{k+1})$, respectively. Dennis et al.proposed the strategy such that the spectrum of the sized matrix $\beta_k A_k$ overlaps that of the second part of the Hessian matrix in the direction of s_k. They obtained the factor (2.9) by using the relation

$$(3.25) \quad | [s_k \{\sum_{i=1}^m f_i^{k+1} \nabla^2 f_i^{k+1}\} s_k / s_k^T s_k][s_k^T(\beta_k A_k) s_k / s_k^T s_k]^{-1} | \fallingdotseq$$
$$| (s_k^T v_k)/\{s_k^T(\beta_k A_k) s_k\} | = 1,$$

where the vector v_k is defined in (2.3). The structured quasi-Newton methods with the sizing factors (2.7) and (2.9) are reasonable in the sense that if the function f_{k+1} becomes zero, then $v_k = 0$ and $\beta_k = 0$, so the new matrix A_{k+1} also becomes zero. This fact is based on using the condition (2.3).

Now we can use the above mentioned factors for our factorized versions. Then we have the following updates:

(i)the sized BFGS-type update

$$(3.26)\quad L_{k+1} = \beta_k L_k + ((\beta_k L_k + J_{k+1})s_k / s_k^T B_k^{\#} s_k)((s_k^T B_k^{\#} s_k / s_k^T z_k)^{1/2} z_k - B_k^{\#} s_k)^T,$$

(ii)the sized DFP-type update

$$(3.27)\quad L_{k+1} = \beta_k L_k + (\beta_k L_k + J_{k+1})((s_k^T z_k / z_k^T (B_k^{\#})^{-1} z_k)^{1/2} (B_k^{\#})^{-1} z_k - s_k)(z_k / s_k^T z_k)^T,$$

where z_k is given by (3.4), the factor β_k is the Biggs' sizing factor (2.7) and the matrix $B_k^{\#}$ is rewritten as

$$(3.28)\quad B_k^{\#} = (\beta_k L_k + J_{k+1})^T (\beta_k L_k + J_{k+1}).$$

It is reasonable to use the above because if the function f_{k+1} becomes zero, then the new matrix L_{k+1} becomes zero, so we have the Gauss-Newton direction at the (k+1)-th iteration. Since the DGW's sizing factor (2.9) contains the matrix A_k, we can not use it directly. However, for the factorized version, the strategy similar to the DGW's one can be considered. The factor β_k should be chosen such that the matrix

$$(3.29)\quad (\beta_k L_k)^T (\beta_k L_k) + (\beta_k L_k)^T J_{k+1} + J_{k+1}^T (\beta_k L_k)$$

has the same spectrum as that of the second part of the Hessian matrix in the direction of s_k. So we have the following relation

$$(3.30)\quad |s_k^T v_k| / s_k^T [(\beta_k L_k)^T (\beta_k L_k) + (\beta_k L_k)^T J_{k+1} + J_{k+1}^T (\beta_k L_k)] s_k = 1,$$

which yields

$$(3.31)\quad \beta_k = \{-(L_k s_k)^T J_{k+1} s_k + \mathrm{sgn}((L_k s_k)^T J_{k+1} s_k)\ \xi_k^{1/2}\} / \|L_k s_k\|^2,$$

where

$$(3.32)\quad \xi_k = ((L_k s_k)^T J_{k+1} s_k)^2 + \|L_k s_k\|^2 |s_k^T v_k|$$

and the symbol $\mathrm{sgn}(\zeta)$ denotes the sign of ζ. In practice, it is also reasonable to use the sizing factor

$$(3.33)\quad \beta_k = \min\{|-(L_k s_k)^T J_{k+1} s_k + \mathrm{sgn}((L_k s_k)^T J_{k+1} s_k)\ \xi_k^{1/2}| / \|L_k s_k\|^2,\ 1\}$$

in the above sense. Though the idea seems interesting, its computational cost is more expensive than that of (2.7).

3.4 New Algorithms

Now we present the two kinds of structured quasi-Newton methods. First, we show the algorithm described in the subsection 3.1 as follows:

(FACNLS Algorithm)

Starting with a point $x_1 \in R^n$ and an $m \times n$ matrix L_1, the algorithm proceeds, for $k = 1, 2, \ldots$, as follows:

Step 1. Having x_k and L_k, find the search direction d_k by solving the linear system of equations

$$(L_k + J_k)^T(L_k + J_k)\, d_k = -J_k^T f_k. \tag{3.34}$$

Step 2. Choose a steplength α_k by a suitable line search algorithm.

Step 3. Set $x_{k+1} = x_k + \alpha_k d_k$.

Step 4. If the new point satisfies the convergence criterion, then stop; otherwise, go to Step 5.

Step 5. Construct L_{k+1} by using the updating formula (3.26) or (3.27).

Next we present the algorithm described in the subsection 3.2:

(INVNLS Algorithm)

Starting with a point $x_1 \in R^n$ and an $n \times n$ symmetric positive definite matrix H_1, the algorithm proceeds, for $k = 1, 2, \ldots$, as follows:

Step 1. Having x_k and H_k, calculate the search direction d_k by

$$d_k = -H_k J_k^T f_k, \tag{3.35}$$

Step 2. Choose a steplength α_k by a suitable line search algorithm.

Step 3. Set $x_{k+1} = x_k + \alpha_k d_k$.

Step 4. If the new point satisfies the convergence criterion, then stop; otherwise, go to Step 5.

Step 5. Construct H_{k+1} by using the updating formula (3.22) or (3.23).

If z_k is given by (3.4), the information of the second part of the Hessian matrix $\nabla^2 F(x)$ is contained only in the secant condition. Further, in the case where $f_{k+1} = 0$, we have $H_{k+1}s_k = (J_{k+1}^T J_{k+1})^{-1}s_k$.

4. Local and Q-Superlinear Convergence of FACNLS Algorithm

We prove the local and q-superlinear convergence of FACNLS algorithms. Our proof is based on the bounded deterioration theorem by Broyden et al. [3]. Let D be the open convex subset of R^n which contains the minimum point x^*. We make the following two assumptions throughout this section:

(A1) There exist positive constants ξ_1, ξ_2 and p such that

$$\|\nabla^2 F(u)-\nabla^2 F(x^*)\| \leqq \xi_1 \|u-x^*\|^p \quad \text{for any } u \text{ in } D, \tag{4.1}$$

and

$$\|J(u_1)-J(u_2)\| \leqq \xi_2 \|u_1-u_2\|^p \quad \text{for any } u_1 \text{ and } u_2 \text{ in } D. \tag{4.2}$$

(A2) $\nabla^2 F$ is symmetric positive definite at x^*.

By using the above assumptions, we have

$$\|\nabla F(u_1)-\nabla F(u_2)-\nabla^2 F(x^*)(u_1-u_2)\| \leqq \xi_1 \max(\|u_1-x^*\|, \|u_2-x^*\|)^p \|u_1-u_2\| \tag{4.3}$$

for any u_1 and u_2 in D, and

$$\|J(u)\| \leqq \xi_2 \|u-x^*\|^p + \|J(x^*)\| \quad \text{for any } u \text{ in } D. \tag{4.4}$$

At first, we show the local and q-superlinear convergence of our algorithm with the DFP-type update (3.14). The following is the key lemma, which is shown by Broyden, Dennis and More [3, Lemma 5.2].

Lemma 1. Let M be an n×n nonsingular symmetric matrix which satisfies

$$\|Mc-M^{-1}a\| \leqq \gamma \|M^{-1}a\|$$

for some $\gamma \in [0,1/3]$ and vectors a and c in R^n with $c^T a \neq 0$. Let X be an n×n symmetric matrix, b any vector in R^n and define Y by

$$Y = X + \{(b-Xa)c^T + c(b-Xa)^T\}/c^Ta - \{a^T(b-Xa)/(c^Ta)^2\}cc^T.$$

Then, for any n×n symmetric G,

$$\|Y-G\|_{F,M} \leqq [(1-\alpha\theta^2)^{1/2} + (5/2)\|Mc-M^{-1}a\|/\{(1-\gamma)\|M^{-1}a\|\}]\|X-G\|_{F,M} + 2(1+2n^{1/2})\|M\|_F\|b-Ga\|/\|M^{-1}a\|,$$

where $\alpha = (1-2\gamma)/(1-\gamma^2) \in [3/8, 1]$, $\theta = 0$ if $X = G$, and $\theta = \|M(X-G)a\|/(\|X-G\|_{F,M}\|M^{-1}a\|)$ if $X \neq G$.

Let M be $\nabla^2 F(x^*)^{-1/2}$. By the equivalence of norms for any n×n matrix

C, there exists a positive constant η such that

(4.5) $\quad \|C\| \leqq \eta \|C\|_{F,M}$.

For each iteration, set

(4.6) $\quad \sigma_k = \max(\|x_k - x^*\|, \|x_{k+1} - x^*\|)$.

Theorem 2. Suppose that the assumptions (A1) and (A2) are satisfied. Let the matrix L_k be updated by the DFP-type formula (3.14), where z_k is given by (3.3) or (3.4). Let the sequence $\{x_k\}$ be generated by

(4.7) $\quad x_{k+1} = x_k - ((L_k + J_k)^T(L_k + J_k))^{-1} J_k^T f_k$.

Then, for any $r \in (0, 1)$, there exist positive constants $\varepsilon(r)$ and $\delta(r)$ such that if $\|x_1 - x^*\| \leqq \varepsilon(r)$ and $\|(L_1+J_1)^T(L_1+J_1) - \nabla^2 F(x^*)\|_{F,M} \leqq \delta(r)$, the sequence $\{x_k\}$ generated by (4.7) is well defined and converges q-linearly to x^* with

(4.8) $\quad \|x_{k+1} - x^*\| \leqq r \|x_k - x^*\|, \quad k \geqq 1.$

Further, $\{\|(L_k+J_k)^T(L_k+J_k)\|\}$, $\{\|((L_k+J_k)^T(L_k+J_k))^{-1}\|\}$, $\{\|(L_k+J_{k+1})^T(L_k+J_{k+1})\|\}$ and $\{\|((L_k+J_{k+1})^T(L_k+J_{k+1}))^{-1}\|\}$ are uniformly bounded.

Proof. For given $r \in (0, 1)$, choose δ such that

(4.9) $\quad 0 < \delta \leqq r/(2\eta\zeta_3)$,

and choose ε so small that

(4.10) $\quad \varepsilon \leqq 1$,

(4.11) $\quad \varepsilon^p \leqq (r/\zeta_3 - 2\eta\delta)/\xi_1$,

(4.12) $\quad 2^{p+1}(1+r)\eta n^2 \xi_2 \|M\|^4 (\zeta_1^{1/2} + 2\zeta_2)\varepsilon^p \leqq 9/10$,

(4.13) $\quad \varepsilon^p \leqq 1/(3\|\nabla^2 F(x^*)^{-1/2}\|^2 \zeta_4)$

and

(4.14) $\quad (2\mu_1\delta + \mu_2)\varepsilon^p/(1-r^p) \leqq \delta$

with

(4.15) $\quad \zeta_1 = 2\eta\delta + \|\nabla^2 F(x^*)\|$,

(4.16) $\quad \zeta_2 = \xi_2 + \|J(x^*)\|$,

(4.17) $\zeta_3 = (1+r)\|\nabla^2 F(x^*)^{-1}\|$,

(4.18) $\zeta_4 = \xi_1 + 2^p(p+2)\xi_2\zeta_2/(p+1)$,

(4.19) $\mu_1 = (15/4)\zeta_4\|M\|^2$,

(4.20) $\mu_2=(n\|M\|)^2\xi_2 2^{p+1}(1+\mu_1)(\zeta_1^{1/2}+2\zeta_2)+2(1+2n^{1/2})\zeta_4\|M\|_F\|M\|$.

Set

(4.21) $N_1 = \{ x \in R^n \mid \|x - x^*\| \leqq \varepsilon \}$

and

(4.22) $N_2 = \{ B \in R^{n\times n} \mid \|B - \nabla^2 F(x^*)\|_{F,M} \leqq 2\delta \}$.

Now we prove, by using the mathematical induction, that the following expressions hold for all $k \geqq 1$:

(E1;k) $\|B_k\| \leqq \zeta_1$ and $B_k \in N_2$,

(E2;k) $\|L_k\| \leqq \zeta_1^{1/2} + \zeta_2$,

(E3;k) $\|B_k^{-1}\| \leqq \zeta_3$,

(E4;k) $\|x_{k+1} - x^*\| \leqq r\|x_k - x^*\|$ and $x_{k+1} \in N_1$,

(E5;k) $\|B_k^{\#}\| \leqq 2^{p+1}\eta n^2\xi_2\varepsilon^p\|M\|^2(\zeta_1^{1/2} + 2\zeta_2) + \zeta_1$,

(E6;k) $\|(B_k^{\#})^{-1}\| \leqq 10\zeta_3$,

(E7;k) The matrix B_{k+1} in (3.15) is well defined and

$$\|B_{k+1} - \nabla^2 F(x^*)\|_{F,M} \leqq \{(1-3\theta_k^2/8)^{1/2} + \mu_1\sigma_k^p\}\|B_k - \nabla^2 F(x^*)\|_{F,M} + \mu_2\sigma_k^p,$$

$$\leqq (1 + \mu_1\sigma_k^p)\|B_k - \nabla^2 F(x^*)\|_{F,M} + \mu_2\sigma_k^p,$$

where σ_k, μ_1 and μ_2 are defined in (4.6),(4.19) and (4.20), respectively, and θ_k is given by

(4.23) $\theta_k = 0$ if $B_k^{\#} = \nabla^2 F(x^*)$

and

(4.24) $\theta_k = \|M(B_k^{\#} - \nabla^2 F(x^*))s_k\| / (\|B_k^{\#} - \nabla^2 F(x^*)\|_{F,M}\|M^{-1}s_k\|)$, otherwise.

First, we consider the case of $k = 1$.

(E1;1) Since

(4.25) $\|B_1 - \nabla^2 F(x^*)\|_{F,M} \leqq \delta \leqq 2\delta$,

it is clear that $B_1 \in N_2$ and

$$\|B_1\| \leq \eta \|B_1 - \nabla^2 F(x^*)\|_{F,M} + \|\nabla^2 F(x^*)\| \leq 2\eta\delta + \|\nabla^2 F(x^*)\|. \tag{4.26}$$

(E2;1) By (4.4) and (4.26), we have

$$\|L_1\| \leq \|L_1 + J(x_1)\| + \|J(x_1)\| \leq \|B_1\|^{1/2} + \|J(x_1)\|$$

$$\leq (2\eta\delta + \|\nabla^2 F(x^*)\|)^{1/2} + \xi_2 \varepsilon^p + \|J(x^*)\| \tag{4.27}$$

(E3;1) It follows from (4.9) and (4.25) that

$$\|\nabla^2 F(x^*)^{-1}\| \|B_1 - \nabla^2 F(x^*)\| \leq 2\eta\delta \|\nabla^2 F(x^*)^{-1}\|$$

$$\leq r/(1+r) < 1.$$

By Banach Perturbation Lemma, B_1 is nonsingular, so B_1 is positive definite, and we have

$$\|B_1^{-1}\| \leq (1+r)\|\nabla^2 F(x^*)^{-1}\|. \tag{4.28}$$

(E4;1) By (4.3), (4.11), (4.25) and (4.28), we have

$$\|x_2 - x^*\| = \|(x_1 - B_1^{-1}\nabla F(x_1)) - x^*\|$$

$$\leq \|B_1^{-1}\| \|\nabla F(x_1) - \nabla F(x^*) - \nabla^2 F(x^*)(x_1 - x^*)\|$$

$$+ \|B_1^{-1}\| \|B_1 - \nabla^2 F(x^*)\| \|x_1 - x^*\|$$

$$\leq (1+r)(2\eta\delta + \xi_1 \varepsilon^p)\|\nabla^2 F(x^*)^{-1}\| \|x_1 - x^*\|$$

$$\leq r \|x_1 - x^*\| < \|x_1 - x^*\| \leq \varepsilon.$$

Thus $x_2 \in N_1$.

(E5;1) Since $J(x_2)$ is available, the matrix $B_1^{\#}$ is well defined. Thus we have

$$\|B_1^{\#} - B_1\|_{F,M} \leq \|(L_1+J_2)^T(J_2 - J_1)\|_{F,M} + \|(J_2 - J_1)^T(L_1 + J_1)\|_{F,M}$$

$$\leq \|M\|_F^2 (\|L_1 + J_2\|_F \|J_2 - J_1\|_F + \|J_2 - J_1\|_F \|L_1 + J_1\|_F)$$

$$\leq n^2 \|M\|^2 (2\|L_1\| + \|J_1\| + \|J_2\|) \|J_2 - J_1\|$$

$$\leq 2^{p+1} n^2 \|M\|^2 \xi_2 \{(2\eta\delta + \|\nabla^2 F(x^*)\|)^{1/2} + 2(\xi_2 \varepsilon^p + \|J(x^*)\|)\} \sigma_1^p. \tag{4.29}$$

Therefore, we have

$$\|B_1^{\#}\| \leq \|B_1^{\#} - B_1\| + \|B_1\|$$

$$\leq 2^{p+1} n^2 \|M\|^2 \xi_2 \eta \{(2\eta\delta + \|\nabla^2 F(x^*)\|)^{1/2}$$

$$+ 2(\xi_2 \varepsilon^p + \|J(x^*)\|)\} \varepsilon^p + 2\eta\delta + \|\nabla^2 F(x^*)\|.$$

(E6;1) By (4.12), (4.28) and (4.29), we have

$$\|B_1^{-1}\|\,\|B_1^{\#}-B_1\| \leqq 2^{p+1}n^2\|M\|^2\xi_2\eta\{(2\eta\delta + \|\nabla^2F(x^*)\|)^{1/2} + 2(\xi_2\varepsilon^p + \|J(x^*)\|)\}\|\nabla^2F(x^*)^{-1}\|(1+r)\varepsilon^p \leqq 9/10 < 1.$$

Thus, by Banach Purterbation Lemma, the matrix $B_1^{\#}$ is nonsingular, so $B_1^{\#}$ is positive definite, and we have

$$\|(B_1^{\#})^{-1}\| \leqq 10(1+r)\|\nabla^2F(x^*)^{-1}\|.$$

(E7;1) At first, we show that there holds for ζ_4 in (4.18)

(4.30) $$\|z_1-\nabla^2F(x^*)s_1\| \leqq \zeta_4\sigma_1^p\|s_1\|,$$

where z_1 is defined by (3.3) or (3.4).

If $z_1 = y_1$, then it follows directly from (4.3) that

$$\|y_1-\nabla^2F(x^*)s_1\| \leqq \xi_1\sigma_1^p\|s_1\|.$$

Consider the case of $z_1 = v_1 + J_2^TJ_2s_1$. Since $f_2-f_1-J_2s_1 = \int_0^1 J(x_1+ts_1)s_1dt - J_2s_1$, we have, by (4.2),

$$\|z_1-\nabla^2F(x^*)s_1\| \leqq \|y_1-\nabla^2F(x^*)s_1\| + \|J_1^T\|\,\|f_2-f_1-J_2s_1\| + \|(J_2-J_1)^T\|\,\|J_2\|\,\|s_1\|.$$

$$\leqq\{\xi_1+(p+2)\xi_2 2^p(\xi_2\varepsilon^p+\|J(x^*)\|)/(p+1)\}\sigma_1^p\|s_1\|.$$

Next, we show that $s_1^Tz_1 > 0$ if $s_1 \neq 0$. This can be shown by a similar way to the proof of Lemma 4.2(a) in [3]. Since

(4.31) $$\|Mz_1-M^{-1}s_1\| \leqq \|M\|\,\|z_1-\nabla^2F(x^*)s_1\| \leqq \zeta_4\sigma_1^p\|M\|\,\|s_1\| \leqq (1/3)\|M^{-1}s_1\|$$

and

$$s_1^Tz_1 = (M^{-1}s_1)^T(Mz_1-M^{-1}s_1) + (M^{-1}s_1)^T(M^{-1}s_1),$$

we have

$$|s_1^Tz_1 - \|M^{-1}s_1\|^2| \leqq \|M^{-1}s_1\|\,\|Mz_1-M^{-1}s_1\| \leqq (1/3)\|M^{-1}s_1\|^2.$$

Thus

$$(2/3)\|M^{-1}s_1\|^2 \leqq s_1^Tz_1 \leqq (4/3)\|M^{-1}s_1\|^2,$$

which suggests that $s_1 \neq 0$ implies $s_1^T z_1 > 0$.

Since the matrix $B_1^{\#}$ is symmetric positive definite and $s_1^T z_1 > 0$, L_2 is well defined, and is given by

$$L_2 = L_1 + (L_1+J_2)((s_1^T z_1/z_1^T (B_1^{\#})^{-1} z_1)^{1/2} (B_1^{\#})^{-1} z_1 - s_1)(z_1/s_1^T z_1)^T.$$

Setting $B_1 = (L_1+J_1)^T(L_1+J_1)$, $B_2 = (L_2+J_2)^T(L_2+J_2)$ and $B_1^{\#} = (L_1+J_2)^T(L_1+J_2)$ gives

$$B_2 = B_1^{\#} + (1+s_1^T B_1^{\#} s_1/s_1^T z_1) z_1 z_1^T/s_1^T z_1 - (B_1^{\#} s_1 z_1^T + z_1 s_1^T B_1^{\#})/s_1^T z_1.$$

Note that the above corresponds to the DFP update. Let

$$X = B_1^{\#}, \quad Y = B_2 \quad , \quad G = \nabla^2 F(x^*), \quad \gamma = 1/3, \quad a = s_1, \quad b = z_1, \quad c = z_1.$$

Then using Lemma 1 and (4.31), we have the bounded deterioration property for $B_1^{\#}$ and B_2 such that

$$\text{(4.32)} \qquad \| B_2 - \nabla^2 F(x^*) \|_{F,M} \leq \{(1-3\theta_1^2/8)^{1/2} + \tau_1 \sigma_1^p\} \| B_1^{\#} - \nabla^2 F(x^*) \|_{F,M} + \tau_2 \sigma_1^p,$$

where

$$\tau_1 = (15/4) \| M \|^2 \zeta_4, \qquad \tau_2 = 2(1 + 2n^{1/2}) \| M \| \| M \|_F \zeta_4,$$

and θ_1 is given by (4.23) or (4.24). Moreover, by noting that

$$\| B_1^{\#} - \nabla^2 F(x^*) \|_{F,M} \leq \| B_1^{\#} - B_1 \|_{F,M} + \| B_1 - \nabla^2 F(x^*) \|_{F,M}$$

and using (4.32),we obtain the bounded deterioration property for B_1 and B_2.

Assume that the expressions (E1;k) through (E7;k) hold for k = 1,..., m-1. Then we have

$$\| B_{k+1} - \nabla^2 F(x^*) \|_{F,M} - \| B_k - \nabla^2 F(x^*) \|_{F,M} \leq (\mu_1 \| B_k - \nabla^2 F(x^*) \|_{F,M} + \mu_2) \sigma_k^p \leq (2\mu_1 \delta + \mu_2) \sigma_k^p$$

for k = 1,..., m-1, and by summing both sides from k = 1 to m-1, it follows from (4.14) that

$$\| B_m - \nabla^2 F(x^*) \|_{F,M} \leq \| B_1 - \nabla^2 F(x^*) \|_{F,M} + (2\mu_1 \delta + \mu_2) \| x_1 - x^* \| \sum_{j=1}^{m-2} (r^p)^j \leq \delta + (2\mu_1 \delta + \mu_2) \varepsilon^p/(1-r^p) \leq 2\delta,$$

which implies (E1;m). We can prove (E2;m) through (E7;m) by the same way as the case of k = 1.

Therefore, this concludes the induction, which completes the proof. ■

Further, by combining Theorem 3.4 in [5] and (E7;k),

$$\lim_{k\to\infty} \| \nabla^2 F(x^*)^{-1/2} B_k \nabla^2 F(x^*)^{-1/2} - I \|_F \tag{4.33}$$

exists and we have

$$\lim_{k\to\infty} \| (B_k - \nabla^2 F(x^*)) s_k \| / \| s_k \| = 0 \tag{4.34}$$

Thus, by Theorem 2.2 in [5], we obtain the following theorem.

Theorem 3. Suppose that all conditions of Theorem 2 hold. Then the sequence $\{x_k\}$ converges q-superlinearly to x^*, that is,

$$\lim_{k\to\infty} \| x_{k+1} - x^* \| / \| x_k - x^* \| = 0. \tag{4.35}$$

Next consider the BFGS-type update (3.9) of L_k. Let $H_k = B_k^{-1}$ and $H_k^{\#} = (B_k^{\#})^{-1}$. Then the relation between $H_k^{\#}$ and H_{k+1} can be given by

$$H_{k+1} = H_k^{\#} + (1 + z_k^T H_k^{\#} z_k / z_k^T s_k) s_k s_k^T / z_k^T s_k - (H_k^{\#} z_k s_k^T + s_k z_k^T H_k^{\#}) / z_k^T s_k, \tag{4.36}$$

which is the form obtained by performing the interchange $B_k^{\#} \longleftrightarrow H_k^{\#}$ and $s_k \longleftrightarrow z_k$ in (3.15). So we can prove the local and q-superlinear convergence of our algorithm with (3.9) by the same means as the above.

Let M be $\nabla^2 F(x^*)^{1/2}$. By the equivalence of norms for any $n \times n$ matrix C, there exist positive constants η and η' such that

$$(1/\eta') \| C \|_{F,M} \leqq \| C \| \leqq \eta \| C \|_{F,M} . \tag{4.37}$$

Theorem 4. Suppose that the assumptions (A1) and (A2) are satisfied. Let the matrix L_k be updated by the BFGS-type formula (3.9), where z_k is given by (3.3) or (3.4). Let the sequence $\{x_k\}$ be generated by

$$x_{k+1} = x_k - ((L_k + J_k)^T (L_k + J_k))^{-1} J_k^T f_k. \tag{4.38}$$

Then, for any $r \in (0, 1)$, there exist positive constants $\varepsilon(r)$ and $\delta(r)$ such that if $\| x_1 - x^* \| \leqq \varepsilon(r)$ and $\| ((L_1 + J_1)^T (L_1 + J_1))^{-1} - \nabla^2 F(x^*)^{-1} \|_{F,M} \leqq \delta(r)$, the sequence $\{x_k\}$ generated by (4.38) is well defined and converges q-linearly to x^* with

$$\| x_{k+1} - x^* \| \leqq r \| x_k - x^* \|, \quad k \geqq 1. \tag{4.39}$$

Further, $\{\|(L_k+J_k)^T(L_k+J_k)\|\}$, $\{\|((L_k+J_k)^T(L_k+J_k))^{-1}\|\}$, $\{\|(L_k+J_{k+1})^T(L_k+J_{k+1})\|\}$ and $\{\|((L_k+J_{k+1})^T(L_k+J_{k+1}))^{-1}\|\}$ are uniformly bounded.

Proof. For given $r \in (0, 1)$, choose δ such that

$$0 < \delta \leqq r/(2\eta\zeta_3'), \tag{4.40}$$

and choose ε so small that

$$\varepsilon \leqq 1, \tag{4.41}$$

$$\zeta_1'\xi_1\varepsilon^p \leqq r^2/(1+r), \tag{4.42}$$

$$2^{p+1}\xi_2\zeta_1'(2\zeta_2' + \zeta_3'^{1/2})\varepsilon^p \leqq 9/10, \tag{4.43}$$

$$\varepsilon^p \leqq 1/(4\|\nabla^2F(x^*)^{-1}\|\,\|M\|\,\|M^{-1}\|\,\zeta_4') \tag{4.44}$$

and

$$(2\mu_1'\delta + \mu_2')\varepsilon^p/(1-r^p) \leqq \delta \tag{4.45}$$

with

$$\zeta_1' = 2\eta\delta + \|\nabla^2F(x^*)^{-1}\|, \tag{4.46}$$

$$\zeta_2' = \xi_2 + \|J(x^*)\|, \tag{4.47}$$

$$\zeta_3' = (1+r)\|\nabla^2F(x^*)\|, \tag{4.48}$$

$$\zeta_4' = \xi_1 + 2^p(p+2)\xi_2\zeta_2'/(p+1), \tag{4.49}$$

$$\mu_1' = 5\zeta_4'\|M\|\,\|M^{-1}\|\,\|\nabla^2F(x^*)^{-1}\|, \tag{4.50}$$

$$\mu_2' = 10\xi_2 2^{p+1}\eta'(1+\mu_1')\zeta_1'^2(2\zeta_2' + (\zeta_3')^{1/2}) + (8/15)(1 + 2n^{1/2})\mu_1'\|M\|_F\|M^{-1}\|. \tag{4.51}$$

Set

$$N_1' = \{ x \in R^n \mid \|x - x^*\| \leqq \varepsilon \} \tag{4.52}$$

and

$$N_2' = \{ H \in R^{n\times n} \mid \|H - \nabla^2F(x^*)^{-1}\|_{F,M} \leqq 2\delta \}. \tag{4.53}$$

Now we prove, by using the mathematical induction, that the following expressions hold for all $k \geqq 1$:

(E1;k)' $\|H_k\| \leqq \zeta_1'$ and $H_k \in N_2'$,

(E2;k)' $\|B_k\| \leqq \zeta_3'$,

(E3;k)' $\| L_k \| \leqq \zeta_2' + (\zeta_3')^{1/2}$,

(E4;k)' $\| x_{k+1} - x^* \| \leqq r \| x_k - x^* \|$ and $x_{k+1} \in N_1'$,

(E5;k)' $\| B_k^{\#} \| \leqq 2^{p+1} \xi_2 (2\zeta_2' + (\zeta_3')^{1/2}) + \zeta_3'$,

(E6;k)' $\| H_k^{\#} \| \leqq 10 \zeta_1'$,

(E7;k)' H_{k+1} is well defined and

$$\| H_{k+1} - \nabla^2 F(x^*)^{-1} \|_{F,M} \leqq \{(1-3\theta_k^2/8)^{1/2} + \mu_1' \sigma_k^p\} \| H_k - \nabla^2 F(x^*)^{-1} \|_{F,M} + \mu_2' \sigma_k^p,$$

$$\leqq (1 + \mu_1' \sigma_k^p) \| H_k - \nabla^2 F(x^*)^{-1} \|_{F,M} + \mu_2' \sigma_k^p,$$

where σ_k, μ_1' and μ_2' are defined in (4.6),(4.50) and (4.51), respectively, and θ_k is given by

(4.54) $\theta_k = 0$ if $H_k^{\#} = \nabla^2 F(x^*)^{-1}$

and

(4.55) $\theta_k = \| M(H_k^{\#} - \nabla^2 F(x^*)^{-1}) z_k \| / (\| H_k^{\#} - \nabla^2 F(x^*)^{-1} \|_{F,M} \| M^{-1} z_k \|)$,

otherwise.

First, we consider the case of k = 1.

(E1;1)' Since the proof is very similar to that of Theorem 2, it is omitted.

(E2;1)' It follows from (4.40) that

$$\| \nabla^2 F(x^*) \| \| H_1 - \nabla^2 F(x^*)^{-1} \| \leqq 2\eta\delta \| \nabla^2 F(x^*) \| \leqq r/(1+r) < 1.$$

By Banach Perturbation Lemma, we have

$$\| B_1 \| = \| H_1^{-1} \| \leqq (1+r) \| \nabla^2 F(x^*) \|.$$

Since the proofs of (E3;1)' through (E6;1)' are very similar to those of Theorem 2, they are omitted.

(E7;1)' By the same way as the proof of Theorem 2, we can show that

(4.56) $\| z_1 - \nabla^2 F(x^*) s_1 \| \leqq \zeta_4' \sigma_1^p \| s_1 \|$

for ζ_4' in (4.49), where z_1 is defined by (3.3) or (3.4), and that $s_1^T z_1 > 0$ if $s_1 \neq 0$. Using Lemma 1, we have the bounded deterioration property for $H_1^{\#}$ and H_2 such that

(4.57) $\| H_2 - \nabla^2 F(x^*)^{-1} \|_{F,M} \leqq \{(1-3\theta_1^2/8)^{1/2}$

$$+ \tau_1' \sigma_1^p\} \| H_1^{\#} - \nabla^2 F(x^*)^{-1} \|_{F,M} + \tau_2' \sigma_1^p,$$

where

$$\tau_1' = 5 \| M \| \| M^{-1} \| \| \nabla^2 F(x^*)^{-1} \| \zeta_4',$$

$$\tau_2' = (8/3)(1+2n^{1/2}) \| M^{-1} \|^2 \| M \| \| M \|_F \| \nabla^2 F(x^*)^{-1} \| \zeta_4',$$

and θ_1 is given by (4.54) or (4.55).

Moreover, by (E1;1)' and (E6;1)', we have

$$\| H_1^{\#} - H_1 \|_{F,M} = \| H_1^{\#}(B_1^{\#} - B_1) H_1 \|_{F,M} \leq \eta' \| H_1^{\#} \| \| H_1 \| \| B_1^{\#} - B_1 \|$$

$$\leq 10 \eta' (\zeta_1')^2 \, 2^{p+1} \xi_2 (2\zeta_2' + (\zeta_3')^{1/2}) \sigma_1^p$$

and

$$\| H_1^{\#} - \nabla^2 F(x^*)^{-1} \|_{F,M} \leq \| H_1^{\#} - H_1 \|_{F,M} + \| H_1 - \nabla^2 F(x^*)^{-1} \|_{F,M}.$$

Therefore, we obtain the bounded deterioration property for H_1 and H_2. Since the remainder of the proof is very similar to that of Theorem 2, it is omitted. ■

Moreover, by a similar way to the proof of Theorem 3.4 in [5],

(4.58) $$\lim_{k \to \infty} \| \nabla^2 F(x^*)^{1/2} H_k \nabla^2 F(x^*)^{1/2} - I \|_F$$

exists and we have

(4.59) $$\lim_{k \to \infty} \| (H_k - \nabla^2 F(x^*)^{-1}) z_k \| / \| z_k \| = 0.$$

Finally, we show that the above is at least the sufficient condition for the superlinear convergence of the algorithm.

Theorem 5. Suppose that the same assumptions of Theorem 4 hold. Let the sequence $\{x_k\}$ be generated by

(4.60) $$x_{k+1} = x_k - H_k \nabla F(x_k).$$

Then $\{x_k\}$ converges q-superlinearly to x^*.

Proof. For the case of $z_k = y_k$, it is proved in [5, p.559] that (4.59) implies the q-superlinear convergence. So we only consider the DGW secant condition (3.4). Since

$$z_k = y_k - J_k^T (f_{k+1} - f_k - J_{k+1} s_k) + (J_{k+1} - J_k)^T J_{k+1} s_k,$$

we have

$$\| H_k \nabla F(x_{k+1}) \| \leqq \| (H_k - \nabla^2 F(x^*)^{-1}) z_k \| + \| \nabla^2 F(x^*)^{-1} \| \| y_k - \nabla^2 F(x^*) s_k \|$$
$$+ \| H_k - \nabla^2 F(x^*)^{-1} \| \| J_k \| \| f_{k+1} - f_k - J_{k+1} s_k \|$$
$$+ \| H_k - \nabla^2 F(x^*)^{-1} \| \| J_{k+1} - J_k \| \| J_{k+1} \| \| s_k \| .$$

Using (4.2), (4.4), (4.56), (4.58), (4.59) and (E1;k)', a similar way to [5, p.559] yields

(4.61) $$\lim_{k \to \infty} \| \nabla F(x_{k+1}) \| / \| x_{k+1} - x_k \| = 0$$

and

(4.62) $$\lim_{k \to \infty} \| x_{k+1} - x^* \| / \| x_k - x^* \| = 0.$$

Thus the proof is complete. ■

5. Computational Experiments

Computational experiments were performed to compare the factorized versions proposed in this paper with the Gauss-Newton method and the structured quasi-Newton methods from the viewpoint of the number of iterations and the number of the objective function evaluations.

The numerical calculations were carried out in double precision arithmetic on a NEC PC-9801VX personal computer, and the program is coded in FORTRAN 77. For all the methods, the initial matrices A_1 and L_1 are set to zero matrices, and H_1 is set to the unit matrix. The iterative process is terminated

(1) if $\| f(x_k) \|_\infty \leqq \max(TOL1, \varepsilon)$,

or

(2) if $| e_j^T J(x_{k+1})^T f(x_{k+1}) | \leqq \max(TOL2, \varepsilon) \| f(x_{k+1}) \| \| J(x_{k+1}) e_j \|$ for $j=1,\ldots,n$ and $\| x_{k+1} - x_k \|_\infty \leqq \max(TOL3, \varepsilon) \max(\| x_{k+1} \|_\infty, 1.0)$, where e_j denotes the j-th column of the unit matrix,

or

(3) if the number of iterations exceeds the prescribed limit (ITMAX),

or

(4) if the number of function evaluations exceeds the prescribed limit (NFEMAX),

where $\| \cdot \|_\infty$ denotes the maximum norm and ε is a machine epsilon.

Further, the Jacobian matrix is evaluated by the forward difference approximation and the bisection line search method with Armijo's rule

$$F(x_k + \alpha_k d_k) \leqq F(x_k) + 0.1\alpha_k \nabla F(x_k)^T d_k$$

is employed.

In the experiments, we set TOL1 = TOL2 = TOL3 = 10^{-4}, ITMAX = 500 and NFEMAX = 2000. Since the sized DFP-type update (3.27) includes the inverse matrix of $B_k^{\#}$, we used only the sized BFGS-type update (3.26) in the FACNLS methods. The test functions to be minimized are listed as follows [2]:

Problem 1.(Powell)

$F = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4$,

Starting point=(3, -1, 0, 1), Minimum point=(0, 0, 0, 0), Value = 0,

Problem 2.(Freudenstein and Roth)

$F = (-13 + x_1 - 2x_2 + 5x_2^2 - x_2^3)^2 + (-29 + x_1 - 14x_2 + x_2^2 + x_2^3)^2$,

Starting point=(15,-2) or (6,6), Local minimum point = (11.4128, -0.89681), Value = 48.98425; Global minimum point = (5,4), Value = 0,

Problem 3.(Kowalik)

$F = \Sigma_{j=1}^{11} (a_j - x_1(u_j^2 + x_2 u_j)/(u_j^2 + x_3 u_j + x_4))^2$,

The data a_j and u_j are given in [2],

Starting point = (0.25, 0.39, 0.415, 0.39),

Minimum point =(0.19281,0.19128,0.12306,0.13606), Value=3.075 x 10^{-4}

Problem 4.(Jennrich)

$F = \Sigma_{j=1}^{10} (a_j - (\exp(jx_1) + \exp(jx_2)))^2$, where $a_j = 2 + 2j$,

Starting point=(0.3,0.4), Minimum point=(0.25783,0.25783), Value=124.36

Problem 5.(Osborne)

$F = \Sigma_{j=1}^{33} (a_j - (x_1 + x_2\exp(-x_4 t_j) + x_3\exp(-x_5 t_j)))^2$,

The data a_j are given in [2] and $t_j = 10(j-1)$,

Starting point = (0.5, 1.5, -1, 0.01, 0.02),

Minimum point = (0.3754, 1.9358, -1.4647, 0.01287, 0.02212),

Minimum Value = 0.546 x 10^{-4}

The computational results are summarized in Tables 1 through 6. In each table, we use the following symbols;

GN : the Gauss-Newton method,
Biggs: the Bartholomew-Biggs update (2.6),
DGW : the Dennis, Gay and Welsch update (2.8),
FAC0 : the FACNLS algorithm with (3.3) and (3.9),
FAC1 : the FACNLS algorithm with (3.4) and (3.9),
FAC2 : the FACNLS algorithm with (2.7), (3.4) and (3.26),
FAC3 : the FACNLS algorithm with (3.33), (3.4) and (3.26),
INV0 : the INVNLS algorithm with (3.4) and (3.22),
INV1 : the INVNLS algorithm with (3.4) and (3.23),
BFGS : the standard BFGS method,
DFP : the standard DFP method,
IT : the number of iterations for convergence,
EV : the number of the objective function evaluations,
FV : the obtained final function value,
* : the method had failed to converge in the specified number of the objective function evaluations.

Table 1. Results for Problem 1

	IT	EV	FV
GN	9	50	2.3×10^{-9}
Biggs	14	75	7.0×10^{-9}
DGW	14	75	6.6×10^{-9}
FAC0	20	105	6.2×10^{-9}
FAC1	14	75	4.8×10^{-9}
FAC2	14	75	5.4×10^{-9}
FAC3	14	75	5.1×10^{-9}
INV0	27	158	3.5×10^{-9}
INV1	77	419	1.5×10^{-8}
BFGS	32	184	1.0×10^{-8}
DFP	97	513	7.6×10^{-9}

Table 2. Results for Problem 2 with (15, -2)

	IT	EV	FV
GN	105	*	58.02
Biggs	6	21	48.98
DGW	6	21	48.98
FAC0	9	33	48.98
FAC1	7	31	48.98
FAC2	7	31	48.98
FAC3	11	102	48.98
INV0	8	36	48.98
INV1	7	33	48.98
BFGS	9	39	48.98
DFP	10	42	48.98

Table 3. Results for Problem 2 with (6,6)

	IT	EV	FV
GN	5	18	2.8×10^{-17}
Biggs	6	21	8.7×10^{-15}
DGW	6	21	8.1×10^{-15}
FAC0	9	30	8.5×10^{-13}
FAC1	6	21	6.5×10^{-9}
FAC2	6	21	6.2×10^{-15}
FAC3	6	21	6.2×10^{-15}
INV0	92	*	730.41
INV1	92	*	730.41
BFGS	13	57	48.98 ($)
DFP	44	172	48.98 ($)

($)The local minimum is obtained.

Table 4. Results for Problem 3

	IT	EV	FV
GN	19	103	3.075×10^{-4}
Biggs	10	59	3.075×10^{-4}
DGW	12	70	3.075×10^{-4}
FAC0	14	94	3.075×10^{-4}
FAC1	11	67	3.075×10^{-4}
FAC2	10	62	3.075×10^{-4}
FAC3	11	69	3.075×10^{-4}
INV0	29	152	3.075×10^{-4}
INV1	399	*	3.711×10^{-4}
BFGS	32	167	3.075×10^{-4}
DFP	399	*	3.772×10^{-4}

Table 5. Results for Problem 4

	IT	EV	FV
GN	138	*	2.7×10^{3}
Biggs	9	32	124.36
DGW	9	32	124.36
FAC0	10	70	124.36
FAC1	9	54	124.36
FAC2	10	57	124.36
FAC3	15	178	124.36
INV0	11	69	124.36
INV1	11	67	124.36
BFGS	13	74	124.36
DFP	30	129	124.36

Table 6. Results for Problem 5

	IT	EV	FV
GN	6	44	5.465×10^{-5}
Biggs	27	172	5.465×10^{-5}
DGW	21	148	5.465×10^{-5}
FAC0	43	274	5.465×10^{-5}
FAC1	26	190	5.465×10^{-5}
FAC2	18	128	5.465×10^{-5}
FAC3	16	119	5.465×10^{-5}
INV0	93	*	6.560×10^{-5}
INV1	329	*	6.580×10^{-5}
BFGS	51	350	5.465×10^{-5}
DFP	172	*	6.019×10^{-5}

From these tables, we can see that the Gauss-Newton method performed very well for the zero or small residual problems (Tables 1 and 6), but did not necessarily well for the large residual problems (Tables 2 and 5). For all the problems, the FACNLS methods except FAC0, the structured quasi-Newton methods with the Biggs and the DGW updates performed well and were numerically stable. These numerical results suggest that the FACNLS methods are comparable with the structured quasi-Newton methods with the Biggs and the DGW updates. However, the INVNLS methods, the standard BFGS and the standard DFP methods did not perform well compared with the other methods.

6. Concluding Remarks

This paper has been concerned with the iterative methods based on the structured quasi-Newton methods for nonlinear least squares problems. Our idea is to compute the search direction by solving the linear system of equations (3.1). This enables us to obtain descent search directions for the objective function.

We proposed the FACNLS and the INVNLS algorithms. However, since the information of the second part $\sum_{j=1}^{m} f_j(x)\nabla^2 f_j(x)$ in (1.5) is contained only in the secant condition, the INVNLS methods did not perform well compared with the other structured quasi-Newton methods in our experiments.
We recommend the FACNLS algorithm with the sized BFGS-type update (3.26) for practical computations. In addition, the FAC1 seems numerically stable though it does not employ a sizing technique.

References

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

[2]J.T.Betts: Solving the nonlinear least square problem - Application of a general method, Journal of Optimization Theory and Applications, Vol.18, No.4, pp.469-483 (1976).

[3]C.G.Broyden, J.E.Dennis,Jr. and J.J.More: On the local and superlinear convergence of quasi-Newton methods, Journal of Institute of Mathematics and its Applications, Vol.12, pp.223-245 (1973).

[4]J.E.Dennis, Jr.: Some computational techniques for the nonlinear least squares problem, in "Numerical Solution of Systems of Nonlinear Algebraic Equations", G.D.Byrne and C.A.Hall (eds.), Academic Press, New York, pp.157-183 (1973).

[5]J.E.Dennis,Jr. and J.J.More: A characterization of superlinear convergence and its application to quasi-Newton methods, Mathematics of Computation, Vol.28, No.126, pp.549-560 (1974).

[6]J.E.Dennis, Jr. and R.B.Schnabel: A new derivation of symmetric positive definite secant updates, in "Nonlinear Programming 4", O.L.Mangasarian, R.R.Meyer and S.M.Robinson (eds.), Academic Press, New York, pp.167-199, (1981).

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

[8]J.E.Dennis, Jr., D.M.Gay and R.E.Welsch: An adaptive nonlinear least squares algorithm, ACM Transactions on Mathematical Software, Vol.7, No.3, pp.348-368 (1981).

[9]R.R.Meyer:Theoretical and computational aspects of nonlinear regression, in "Nonlinear Programming",J.B.Rosen,O.L.Mangasarian and K.Ritter(eds.), Academic Press, New York, pp.465-486 (1970).

[10]H.Yabe and T.Takahashi: Structured quasi-Newton methods for nonlinear least squares problems, TRU Mathematics, Vol.24, No.2 (1988) to appear.