Journal of the Operations Research
Society of Japan
Vol. 34, No. 3, September 1991

NUMERICAL COMPARISON AMONG STRUCTURED QUASI-NEWTON METHODS FOR NONLINEAR LEAST SQUARES PROBLEMS

Hiroshi Yabe
Science University of Tokyo

Toshihiko Takahashi
Kajima Corporation

(Received May 8, 1990; Revised March 11, 1991)

Abstract The purpose of this paper is to construct effective algorithms for solving nonlinear least squares problems. These methods are based on the idea of structured quasi-Newton methods, which use the structure of the Hessian matrix of the objective function. In order to obtain a descent search direction of the objective function, we have proposed to approximate the Hessian matrix by the factorized form and the BFGS-like update and DEP-like update have been obtained. Independently of us, Sheng Songbai and Zou Zhihong (SZ) have been studying factorized versions of structured quasi-Newton methods. In this paper, we construct an update by a slight different way from their formulation, in which the SZ update is contained. Further, we apply sizing techniques to the SZ method and propose new sizing factors. Finally, computational experiments are shown in order to compare our factorized versions with the SZ method and investigate the effect of sizing techniques.

1. Introduction

This paper is concerned with numerical methods for finding a point x^* which minimizes a sum of squares of nonlinear functions

$$f(x) = \frac{1}{2}\sum_{j=1}^{m}(r_j(x))^2, \quad m \geq n, \tag{1.1}$$

where $r_j : R^n \to R$ is twice continuously differentiable for $j = 1, \ldots, m$. This type of minimization problems typically occurs in curve fitting.

For general unconstrained minimization problems where the Hessian matrix of second derivatives can be calculated, Newton's method can be used. The method constructs a sequence of vectors $\{x_k\}$ such that

$$x_{k+1} = x_k + \alpha_k d_k, \tag{1.2}$$

where α_k is a scalar steplength and d_k, the direction of search, satisfies the Newton equation

$$\nabla^2 f(x_k)d = -\nabla f(x_k). \tag{1.3}$$

For a sum of squares of nonlinear functions, the gradient vector and Hessian matrix have the special forms, which are given by

$$\nabla f(x) = J(x)^T r(x) \tag{1.4}$$

and

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

respectively, where

$$r(x) = (r_1(x), \ldots, r_m(x))^T$$

and $J(x)$ is the $m \times n$ Jacobian matrix of $r(x)$ whose i-th row is $\nabla r_i(x)^T$, 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, methods have been derived which use only the first derivative information. For example, the Gauss-Newton method and the Levenberg-Marquardt method exploit the special structure of the Hessian matrix and gradient vector. Since these methods neglect the second part of the Hessian matrix of f, they can be expected to perform well when the residuals at x^* are small or each r_i is close to linear. We call these cases the small residual problems. However, when the residuals at x^* are very large and the functions are rather nonlinear, these methods may perform poorly. We call these cases the large residual problems.

Recently, quasi-Newton approximations to only the second part of the Hessian matrix have been developed [5]. We call these strategies the structured quasi-Newton methods. Among these methods, typical methods are the line search descent method and the trust region method. The former has been studied by Al-Baali and Fletcher [1], Bartholomew-Biggs [2], Fletcher and Xu [7]. The latter has been studied by Dennis, Gay and Welsch [6]. In this paper, we only consider the former. In which case, it is desirable to maintain the positive definiteness of the coefficient matrix of the Newton equation, which enables us to obtain descent search directions. Within this framework, factorized versions of structured quasi-Newton methods have been studied independently by Sheng Songbai and Zou Zhihong [9], and Yabe and Takahashi [10].

It seems that Sheng Songbai and Zou Zhihong introduced their updating formula without investigating the existence of such a formula in detail. In this paper, we examine closely its existence and construct an updating formula which contains their update, though we use the same least change secant update as Sheng Songbai et al., and we apply sizing techniques to their method. In Section 2 and Section 3, we review the structured quasi-Newton methods and our factorized versions, respectively. In Section 4, we construct an update which contains the Sheng Songbai and Zou Zhihong (SZ) update. In Section 5, we apply sizing techniques to the SZ update and propose new sizing factors. Further, a factorized algorithm is given in Section 6. Finally, computational experiments are shown in order to compare our methods with the SZ method and investigate the effect of sizing techniques.

Throughout this paper, $\|\bullet\|$ denotes the 2-norm for vectors or matrices. $\|Q\|_F$ denotes the Frobenius norm of a matrix Q and is defined by

$$\|Q\|_F = \sqrt{\mathrm{Trace}(QQ^T)}.$$

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

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 d_k can be computed by solving

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

where $r_k = r(x_k)$, $J_k = J(x_k)$, and the matrix A_k is the k-th approximation to the second part of the Hessian matrix of f. The matrix A_k is updated such that the new matrix A_{k+1} satisfies the secant condition

$$A_{k+1}s_k = y_k - J_{k+1}^T J_{k+1} s_k \tag{2.2}$$

or

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

where

$$s_k = x_{k+1} - x_k, \quad y_k = \nabla f(x_{k+1}) - \nabla f(x_k). \tag{2.4}$$

By using sizing techniques, Bartholomew-Biggs [2], and Dennis, Gay and Welsch (DGW) [6] proposed the robust algorithms for both cases of large and small residual problems. Their updates are as follows:
(i) the Biggs update

$$A_{k+1} = \beta_k A_k + \frac{(v_k - \beta_k A_k s_k)(v_k - \beta_k A_k s_k)^T}{(v_k - \beta_k A_k s_k)^T s_k}, \tag{2.5}$$

$$\beta_k = \frac{r_{k+1}^T r_k}{r_k^T r_k}, \tag{2.6}$$

(ii) the DGW update

$$A_{k+1} = \beta_k A_k + \frac{(v_k - \beta_k A_k s_k) y_k^T + y_k (v_k - \beta_k A_k s_k)^T}{s_k^T y_k} - \frac{s_k^T (v_k - \beta_k A_k s_k)}{(s_k^T y_k)^2} y_k y_k^T, \tag{2.7}$$

$$\beta_k = \min\left(\left| \frac{s_k^T v_k}{s_k^T A_k s_k} \right|, \ 1 \right), \tag{2.8}$$

where β_k is a sizing factor.

3. Factorized Versions of Structured Quasi-Newton Methods

For general quasi-Newton methods, the hereditary positive definiteness property is desirable because a descent search direction for the objective function is obtained. On the other hand, for structured quasi-Newton methods, it is not clear how to construct updating formulae for 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 trust region strategy) [6] and the hybrid method [1], [7].

In [10], we proposed a direct approach which maintains positive definiteness of the coefficient matrix in (2.1). We compute the search direction d_k by solving the linear system of equations

$$(J_k + L_k)^T (J_k + L_k) d = -J_k^T r_k, \tag{3.1}$$

where the matrix L_k is an $m \times n$ correction matrix to the Jacobian matrix such that $L_k^T J_k + J_k^T L_k + L_k^T L_k$ is the k-th approximation to the second part of the Hessian matrix of f. Since the coefficient matrix is expressed by the factorized form, the search direction may be expected to be a descent direction for f. Successful updates for L_k would lead to simplified line search algorithms in contrast to the more complex trust region algorithms.

The secant condition for L_{k+1} is as follows:

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

where

$$z_k = y_k \tag{3.3}$$

or

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

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 there hold

$$h^T L_{k+1} = (z_k - J_{k+1}^T h)^T, \tag{3.5}$$

$$L_{k+1} s_k = h - J_{k+1} s_k \tag{3.6}$$

for some m-dimensional vector h. In order to find L_{k+1} which satisfies the matrix equations (3.5) and (3.6), we use the following lemma [3].

Lemma 1. *The matrix equations*

$$CX = D, \quad XE = G$$

have a common solution X if and only if each equation separately has a solution and

$$CG = DE,$$

where C, D, E and G are given matrices and X is an unknown matrix.

By using the above, the matrix equations (3.5) and (3.6) have a common solution L_{k+1} if and only if each equation separately has a solution and $h^T h = s_k^T z_k$. Yabe and Takahashi [10] proposed two types of updates by using the least change secant update technique in the sense of Dennis and Schnabel [4], under the assumption of $s_k^T z_k > 0$.

Finding the L_{k+1} which minimizes $\|L_{k+1}^T - L_k^T\|_F$ subject to the condition (3.5) and solving for h so that the other condition (3.6) can hold lead to the BFGS-like update:

$$L_{k+1} = L_k + \frac{(J_{k+1} + L_k) s_k}{s_k^T B_k^{\dagger} s_k} \left(\sqrt{\frac{s_k^T B_k^{\dagger} s_k}{s_k^T z_k}}\, z_k - B_k^{\dagger} s_k \right)^T, \tag{3.7}$$

where

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

On the other hand, for given nonsingular matrices W_L and W_R, finding the L_{k+1} which minimizes $\|W_L (L_{k+1} - L_k) W_R\|_F$ subject to the condition (3.6) and solving for h so that the other condition (3.5) can hold lead to the DFP-like update:

$$L_{k+1} = L_k + (J_{k+1} + L_k) \left(\sqrt{\frac{s_k^T z_k}{z_k^T (B_k^{\dagger})^{-1} z_k}}\, (B_k^{\dagger})^{-1} z_k - s_k \right) \left(\frac{z_k}{s_k^T z_k} \right)^T. \tag{3.8}$$

The local and q-superlinear convergence of these methods are proven in [11].

4. An Update Which Contains the SZ Update

Independently of us, Sheng Songbai and Zou Zhihong [9] have been studying factorized versions of structured quasi-Newton methods. They proposed the approximation of $r(x)$ around x_k as follows:

$$m(x_k + d) = r_k + (J_k + L_k)d. \tag{4.1}$$

If $L_k = 0$, then the above is reduced to the Gauss-Newton model. Here the model function to be minimized is $(1/2)\|m(x_k + d)\|^2$, and the search direction d_k is obtained by solving the normal equation

$$(J_k + L_k)^T(J_k + L_k)d = -(J_k + L_k)^T r_k. \tag{4.2}$$

Since this does not correspond to the Newton equation (1.3), they imposed the condition $L_k^T r_k = 0$ on a matrix L_k, in addition to the secant condition (3.2). So the conditions which the matrix L_{k+1} should satisfy are as follows:

$$(J_{k+1} + L_{k+1})^T(J_{k+1} + L_{k+1})s_k = z_k \tag{4.3}$$

and

$$L_{k+1}^T r_{k+1} = 0, \tag{4.4}$$

where

$$z_k = (J_{k+1} - J_k)^T r_{k+1} + J_{k+1}^T J_{k+1} s_k. \tag{4.5}$$

Similar to the discussion in the previous section, the matrix equation (4.3) is equivalent to the matrix equations

$$h^T L_{k+1} = (z_k - J_{k+1}^T h)^T, \tag{4.6}$$

$$L_{k+1} s_k = h - J_{k+1} s_k \tag{4.7}$$

for some m-dimensional vector h. Here the purpose is to find a rectangular matrix L_{k+1} which satisfies the equations (4.4), (4.6) and (4.7). It seems that Sheng Songbai and Zou Zhihong introduced their updating formula without investigating the existence of such a matrix in detail. So, though we use the same least change secant update, we examine closely the existence of L_{k+1} and construct an updating formula which contains their update.

Assume that r_{k+1} and s_k are not zero vectors. It follows from Lemma 1 that the matrix equations (4.4), (4.6) and (4.7) have a common solution L_{k+1} if and only if

$$h^T h = s_k^T z_k > 0 \quad \text{and} \quad r_{k+1}^T h = r_{k+1}^T J_{k+1} s_k \tag{4.8}$$

are satisfied. Thus, under the assumption (4.8), we only find a rectangular matrix L_{k+1} which satisfies (4.4), (4.6) and (4.7). Now we drop the suffix k and replace the suffix $(k+1)$ by '+' for simplicity of notation. For a matrix M, let $R(M)$ denote a space spanned by column vectors of M. Then we can consider the following two cases:

4.1. Case I

When h is contained in $R(r_+)$, h is represented by

$$h = \pm\sqrt{s^T z}\,\frac{r_+}{\|r_+\|}.$$

If $z - J_+^T h \neq 0$, then the matrix equations (4.4) and (4.6) are inconsistent. Otherwise, since (4.6) is equivalent to (4.4), the conditions are reduced to the expressions

$$L_+ s = \pm\sqrt{s^T z}\,\frac{r_+}{\|r_+\|} - J_+ s \quad \text{and} \quad L_+^T r_+ = 0.$$

By Lemma 1, a necessary and sufficient condition that the above has a common solution is that there holds $\pm\sqrt{s^T z}\,\|r_+\| = r_+^T J_+ s$. However, since it does not seem meaningful to construct a matrix L_+ under this condition, we only consider the case where h is not contained in $R(r_+)$, which is discussed in the following subsection.

4.2. Case II

When h is not contained in $R(r_+)$, we can consider a least change secant update following to Sheng Songbai and Zou Zhihong. For any unknown m-dimensional vector h such that $h^T h = s^T z$ and $r_+^T h = r_+^T J_+ s$, minimizing the Frobenius norm $\|L_+^T - L^T\|_F$ with respect to L_+, subject to $L_+^T r_+ = 0$ in (4.4) and $L_+^T h = z - J_+^T h$ in (4.6), we have a unique solution

$$L_+ = PL + \frac{Ph(z - M^T h)^T}{\|Ph\|^2},$$

where

$$M = J_+ + PL \quad \text{and} \quad P = I - \frac{r_+ r_+^T}{r_+^T r_+}. \tag{4.9}$$

By substituting the above into the other condition (4.7), we have

$$\left(1 - \frac{(z - M^T h)^T s}{\|Ph\|^2}\right) h = Ms - \frac{(z - M^T h)^T s (r_+^T h)}{\|Ph\|^2 \|r_+\|^2}\, r_+. \tag{4.10}$$

Then we can further consider two cases:

4.2.1. Case II-1

When Ms is not contained in $R(r_+)$, $1 - (z - M^T h)^T s/\|Ph\|^2 \neq 0$ should be satisfied. In fact, for an m-dimensional vector h such that $(z - M^T h)^T s = \|Ph\|^2$, the left-hand side of (4.10) becomes zero ; on the other hand, the linear independence of Ms and r_+ implies $Ms - (r_+^T h/\|r_+\|^2) r_+ \neq 0$, which is a contradiction. Consequently, it follows from (4.10) that h can be represented by the form

$$h = \tau_1 Ms + \tau_2 r_+, \quad \tau_1 \neq 0.$$

Substituting the above into the expression (4.10), we have

$$\left(1 - \frac{s^T z - \tau_1 s^T M^T M s - \tau_2 s^T J_+^T r_+}{\tau_1^2 \|PMs\|^2}\right)(\tau_1 Ms + \tau_2 r_+)$$
$$= Ms - \frac{(s^T z - \tau_1 s^T M^T M s - \tau_2 s^T J_+^T r_+)(\tau_1 s^T J_+^T r_+ + \tau_2 \|r_+\|^2)}{\tau_1^2 \|PMs\|^2 \|r_+\|^2}\, r_+.$$

By arranging the coefficients of the vectors Ms and r_+, and using the linear independence of Ms and r_+,

$$\frac{s^T z - \tau_1 s^T M^T M s - \tau_2 s^T J_+^T r_+}{\tau_1 \|PMs\|^2} = \tau_1 - 1, \tag{4.11}$$

$$\frac{(s^T z - \tau_1 s^T M^T M s - \tau_2 s^T J_+^T r_+)\, r_+^T J_+ s}{\tau_1 \|PMs\|^2 \|r_+\|^2} = -\tau_2. \tag{4.12}$$

Then we have

$$\tau_2 = \frac{(1 - \tau_1) r_+^T J_+ s}{\|r_+\|^2}.$$

Substituting the above into the expression (4.11) and setting $\tau = \tau_1$, we have the quadratic equation of τ

$$\|PMs\|^2\tau^2 + \left(\|Ms\|^2 - \|PMs\|^2 - \frac{(r_+^T J_+ s)^2}{\|r_+\|^2}\right)\tau + \left(\frac{(r_+^T J_+ s)^2}{\|r_+\|^2} - s^T z\right) = 0,$$

which yields

$$\|PMs\|^2\tau^2 = s^T z - \frac{(r_+^T J_+ s)^2}{\|r_+\|^2}. \tag{4.13}$$

Since it follows from (4.8) that

$$s^T z - \frac{(r_+^T J_+ s)^2}{\|r_+\|^2} = \|Ph\|^2 > 0,$$

the expression (4.13) can be solved with respect to τ, which leads to

$$h = \tau Ms + (1-\tau)\frac{r_+^T J_+ s}{\|r_+\|^2} r_+. \tag{4.14}$$

Then we have

$$L_+ = PL + \frac{Ph(z - M^T h)^T}{\|Ph\|^2}, \tag{4.15}$$

where τ and h are given in (4.13) and (4.14), respectively, which corresponds to the SZ update. Note that Sheng Songbai et al. assumed the positiveness of $s^T z - (r_+^T J_+ s)^2/\|r_+\|^2$, but the necessary and sufficient condition (4.8) certainly guarantees it. When the condition (4.4) cannot be imposed, the above is reduced to the BFGS-like update (3.7).

4.2.2. Case II-2

When Ms is contained in $R(r_+)$, Ms is formed by

$$Ms = \eta r_+, \qquad \eta = \frac{r_+^T Ms}{\|r_+\|^2}. \tag{4.16}$$

Thus, it follows from (4.10) that

$$\|r_+\|^2\{\|Ph\|^2 - (z - M^T h)^T s\}h = \{\eta\|Ph\|^2\|r_+\|^2 - (z - M^T h)^T s(r_+^T h)\}r_+.$$

Using $\|Ph\|^2 = s^T z - (r_+^T h)^2/\|r_+\|^2$ and (4.16), the expression (4.10) can be represented by

$$(r_+^T h)(r_+^T Ms - r_+^T h)h = (s^T z)(r_+^T Ms - r_+^T h)r_+. \tag{4.17}$$

It follows from (4.8) that h satisfies

$$r_+^T h = r_+^T Ms, \tag{4.18}$$

which makes the coefficients of both sides in (4.17) lead to zero. Then, noting that Ms is a particular solution of the equation (4.18) and that the condition $h^T h = s^T z$ should be satisfied, we have a general solution of the equation (4.18) as follows

$$h = Ms + \frac{\sqrt{s^T z - \|Ms\|^2}}{\|Pu\|} Pu,$$

where P is an orthogonal projection matrix (4.9) onto the null space of r_+ and u is an m-dimensional arbitrary vector which is not contained in $R(r_+)$.

4.2.3. Updating Formula

Finally, summarizing (Case II-1) and (Case II-2), we obtain the following update:

$$
\begin{aligned}
L_{k+1} &= P_k L_k + \frac{P_k h_k (z_k - M_k^T h_k)^T}{\|P_k h_k\|^2}, \qquad (4.19)\\
h_k &= \frac{r_{k+1}^T J_{k+1} s_k}{\|r_{k+1}\|^2} r_{k+1} + \rho_k \frac{P_k w_k}{\|P_k w_k\|},\\
P_k &= I - \frac{r_{k+1} r_{k+1}^T}{\|r_{k+1}\|^2},\\
M_k &= J_{k+1} + P_k L_k,\\
\rho_k^2 &= s_k^T z_k - \frac{(r_{k+1}^T J_{k+1} s_k)^2}{\|r_{k+1}\|^2}, \qquad (4.20)
\end{aligned}
$$

where $w_k = M_k s_k$ if $M_k s_k$ is not contained in $R(r_{k+1})$, i.e. $\|P_k M_k s_k\| \neq 0$, otherwise, w_k is chosen to be linearly independent of r_{k+1}.

Noting $\rho_k^2 = \|P_k M_k s_k\|^2 \tau_k^2$ in the expressions (4.13) and (4.20), it is clear that the update (4.19) with $w_k = M_k s_k$ can be reduced to the SZ update (4.15).

5. Sizing techniques of the updating matrix

We know that, for zero residual problems, the matrices A_k and $L_k^T J_k + J_k^T L_k + L_k^T L_k$ 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. Since the quasi-Newton updates do not generate the zero matrix, some remedies must be employed. Among remedies, the sizing of the updating matrices which has been introduced by Bartholomew-Biggs [2] or Dennis et al. [6] seems most promising. The structured quasi-Newton methods with the sizing factors (2.6) and (2.8) may be reasonable in the sense that if the function r_{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 derived by using the secant condition (2.3).

An application of sizing techniques to factorized versions was proposed by Yabe and Takahashi [10]. They derived the following updates:

(i) the sized BFGS-like update

$$
L_{k+1} = \beta_k L_k + \frac{(J_{k+1} + \beta_k L_k) s_k}{s_k^T B_k^{\sharp} s_k} \left(\sqrt{\frac{s_k^T B_k^{\sharp} s_k}{s_k^T z_k}}\, z_k - B_k^{\sharp} s_k \right)^T, \qquad (5.1)
$$

(ii) the sized DFP-like update

$$
L_{k+1} = \beta_k L_k + (J_{k+1} + \beta_k L_k) \left(\sqrt{\frac{s_k^T z_k}{z_k^T (B_k^{\sharp})^{-1} z_k}} (B_k^{\sharp})^{-1} z_k - s_k \right) \left(\frac{z_k}{s_k^T z_k} \right)^T, \qquad (5.2)
$$

where z_k is given by (3.4), β_k is a suitable sizing factor and the matrix $B_k^{\sharp}$ is rewritten as

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

Applying the same technique to the SZ update (4.15), we have

(iii) the sized SZ update

$$
\begin{aligned}
L_{k+1} &= \beta_k P_k L_k + \frac{P_k h_k (z_k - M_k^T h_k)^T}{\|P_k h_k\|^2}, \\
h_k &= (r_{k+1}^T J_{k+1} s_k)(\|r_{k+1}\|^2)^\dagger r_{k+1} + \rho_k \frac{P_k M_k s_k}{\|P_k M_k s_k\|}, \\
P_k &= I - (\|r_{k+1}\|^2)^\dagger r_{k+1} r_{k+1}^T, \\
M_k &= J_{k+1} + \beta_k P_k L_k, \\
\rho_k &= \sqrt{s_k^T z_k - (r_{k+1}^T J_{k+1} s_k)^2 (\|r_{k+1}\|^2)^\dagger},
\end{aligned} \tag{5.3}
$$

where $q^\dagger$ denotes the Moore-Penrose generalized inverse of q.

Note that we can apply Biggs' sizing factor (2.6) to the factorized quasi-Newton updates. On the other hand, since the DGW sizing factor (2.8) contains the matrix A_k, we cannot employ it directly. However, for the factorized versions, a strategy similar to the DGW's one was considered in [11]. The factor β_k should be chosen such that the matrix

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

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

$$\left| \frac{s_k^T v_k}{s_k^T [(\beta_k L_k)^T J_{k+1} + J_{k+1}^T (\beta_k L_k) + (\beta_k L_k)^T (\beta_k L_k)] s_k} \right| = 1,$$

which yields the quadratic equation of β_k and we have the solution

$$\beta_k' = \frac{-(L_k s_k)^T J_{k+1} s_k + \mathrm{sgn}((L_k s_k)^T J_{k+1} s_k)\sqrt{\phi_k}}{\|L_k s_k\|^2}, \tag{5.4}$$

where $\phi_k = ((L_k s_k)^T J_{k+1} s_k)^2 \pm \|L_k s_k\|^2 (s_k^T v_k)$ and the symbol $\mathrm{sgn}(\zeta)$ denotes the sign of ζ.

In [11], we proposed the sizing factor defined by

$$\beta_k = \min(|\beta_k'|,\ 1), \quad \phi_k = ((L_k s_k)^T J_{k+1} s_k)^2 + \|L_k s_k\|^2 |s_k^T v_k|. \tag{5.5}$$

Here, by investigating the signs of $s_k^T v_k$ and ϕ_k, we can obtain two new strategies (a) and (b):

(a) Set $\phi_k = ((L_k s_k)^T J_{k+1} s_k)^2 + \|L_k s_k\|^2 |s_k^T v_k|$. For β_k' in (5.4), we choose

(a-1)

$$\beta_k = \min(|\beta_k'|,\ 1),$$

or

(a-2)

$$\beta_k = \begin{cases} -1 & \text{if} \quad \beta_k' \le -1, \\ \beta_k' & \text{if} \quad -1 < \beta_k' < 1, \\ 1 & \text{if} \quad 1 \le \beta_k'. \end{cases}$$

Note that, in (a-1) , we use the absolute value of β_k', which corresponds to (5.5), and that, in (a-2) , we consider the sign of β_k'.

(b) Set

$$\begin{aligned}\phi_k^1 &= ((L_k s_k)^T J_{k+1} s_k)^2 + \|L_k s_k\|^2 (s_k^T v_k),\\ \phi_k^2 &= ((L_k s_k)^T J_{k+1} s_k)^2 - \|L_k s_k\|^2 (s_k^T v_k)\end{aligned}$$

and

$$\begin{aligned}\beta_k^1 &= \frac{-(L_k s_k)^T J_{k+1} s_k + \mathrm{sgn}((L_k s_k)^T J_{k+1} s_k)\sqrt{\phi_k^1}}{\|L_k s_k\|^2},\\ \beta_k^2 &= \frac{-(L_k s_k)^T J_{k+1} s_k + \mathrm{sgn}((L_k s_k)^T J_{k+1} s_k)\sqrt{\phi_k^2}}{\|L_k s_k\|^2}.\end{aligned}$$

Then we have
(b-1)

$$\beta_k = \begin{cases} \min\{\max(|\beta_k^1|, |\beta_k^2|), 1\} & \text{if} \quad \phi_k^1 \geq 0 \text{ and } \phi_k^2 \geq 0,\\ \min(|\beta_k^1|, 1) & \text{if} \quad \phi_k^1 \geq 0 \text{ and } \phi_k^2 < 0,\\ \min(|\beta_k^2|, 1) & \text{otherwise,}\end{cases}$$

or
(b-2) Set

$$\beta_k' = \begin{cases} \beta_k^1 & \text{if } (\phi_k^1 \geq 0 \text{ and } \phi_k^2 \geq 0 \text{ and } |\beta_k^1| \geq |\beta_k^2|) \text{ or } \text{ if } (\phi_k^1 \geq 0 \text{ and } \phi_k^2 < 0),\\ \beta_k^2 & \text{otherwise.}\end{cases}$$

For this β_k', we choose β_k by the same way as Strategy (a-2).

6. Algorithm

Now we present an algorithm of a new factorized quasi-Newton method.

(FACTNLS method)

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 (3.1).

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 a suitable updating formula for L_k.

It should be noted that, for the SZ method, Step 1 can be rewritten as

Step 1'. Having x_k and L_k, find the search direction d_k by solving the normal equation (4.2).

7. Computational Experiments

Computational experiments were performed to compare our factorized versions with the SZ method from the viewpoint of the number of iterations and the number of vector valued function (i.e. $r(x)$) evaluations. Further, we examined the effect of sizing techniques.

The numerical calculations were carried out in double precision arithmetic on a NEC PC-9801VX personal computer, and the program was coded in FORTRAN 77. The iterative process is terminated

(i) if $\|r(x_k)\|_\infty \leq \max(\text{TOL1}, \varepsilon)$,

or

(ii) if $|e_i^T J(x_{k+1})^T r(x_{k+1})| \leq \max(\text{TOL2}, \varepsilon)\|r(x_{k+1})\| \, \|J(x_{k+1})e_i\|$ for $i = 1, \ldots, n$ and $\|x_{k+1} - x_k\|_\infty \leq \max(\text{TOL3}, \varepsilon)\max(\|x_{k+1}\|_\infty, 1.0)$, where e_i denotes the i-th column of the unit matrix,

or

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

or

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

where $\| \bullet \|_\infty$ denotes the maximum norm and ε is machine epsilon. Further, the Jacobian matrix is evaluated by the forward difference approximation, and the bisection line search method with Armijo's rule

$$f(x_k + \alpha_k d_k) \leq f(x_k) + 0.1\alpha_k \nabla f(x_k)^T d_k \tag{7.1}$$

is employed. In the experiments, we set TOL1 = TOL2 = TOL3 = 10^{-4}, ITMAX = 500 and NFEMAX = 2000. In addition to (2.6), we used the following sizing factor

$$\beta_k = \frac{|r_{k+1}^T r_k|}{\|r_k\|^2}. \tag{7.2}$$

For all the methods, the initial matrix L_1 was set to the zero matrix.

The names, the sizes and the starting points of the test problems given in [6] and [8], together with the abbreviated problem names used in Tables 2–5, are listed in Table 1. In Table 1, (Z), (S) and (L) mean a zero residual problem, a small residual problem, and a large residual problem, respectively.

The computational results are summarized in Tables 2–5. Note that the numbers in Tables 3 and 5 include the number of vector valued function (i.e. $r(x)$) evaluations to evaluate $J(x)$ by the forward difference approximation. In each table, we use the following symbols, where GN means the Gauss-Newton method:

BFGSF-0	:	the FACTNLS method with (3.3) and (3.7),
F-1	:	the FACTNLS method with (3.4) and (3.7),
F-2a	:	the FACTNLS method with (3.4), (5.1) and (7.2),
F-2b	:	the FACTNLS method with (3.4), (5.1) and (2.6),
F-3a	:	the FACTNLS method with (3.4), (5.1) and (a-1),
F-3b	:	the FACTNLS method with (3.4), (5.1) and (a-2),
F-4a	:	the FACTNLS method with (3.4), (5.1) and (b-1),
F-4b	:	the FACTNLS method with (3.4), (5.1) and (b-2),

DFP F-0	:	the FACTNLS method with (3.3) and (3.8),
F-1	:	the FACTNLS method with (3.4) and (3.8),
F-2a	:	the FACTNLS method with (3.4), (5.2) and (7.2),
F-2b	:	the FACTNLS method with (3.4), (5.2) and (2.6),
F-3a	:	the FACTNLS method with (3.4), (5.2) and (a-1),
F-3b	:	the FACTNLS method with (3.4), (5.2) and (a-2),
F-4a	:	the FACTNLS method with (3.4), (5.2) and (b-1),
F-4b	:	the FACTNLS method with (3.4), (5.2) and (b-2),
SZ F-0	:	the FACTNLS method with (3.3) and (4.15),
F-1	:	the FACTNLS method with (3.4) and (4.15),
F-2a	:	the FACTNLS method with (3.4), (5.3) and (7.2),
F-2b	:	the FACTNLS method with (3.4), (5.3) and (2.6),
F-3a	:	the FACTNLS method with (3.4), (5.3) and (a-1),
F-3b	:	the FACTNLS method with (3.4), (5.3) and (a-2),
F-4a	:	the FACTNLS method with (3.4), (5.3) and (b-1),
F-4b	:	the FACTNLS method with (3.4), (5.3) and (b-2),
G-N1	:	if $\|r_k\| \leq 10^{-1}$, then GN is used, otherwise SZF-1,
G-N2	:	if $\|r_k\| \leq 10^{-3}$, then GN is used, otherwise SZF-1,
G-N3	:	if $\|r_k\| \leq 10^{-5}$, then GN is used, otherwise SZF-1,
*	:	the method failed to converge.

For comparison, the results of BFGSF-0, F-1, F-2a, F-3a, DFPF-0, F-1, F-2a and F-3a in Tables 2 and 3 are referred to from [11]. Further, since the comparison of our methods with the Gauss-Newton method, the Biggs method, and the DGW method was shown in [11], we omit those numerical results. In the above, SZG-N1, SZG-N2 and SZG-N3 are original SZ methods, which combine the structured quasi-Newton method and the Gauss-Newton method.

The following can be observed from Tables 2 and 3. In our factorized methods, BFGSF-0 and DFPF-0 did not perform well for all the problems, and the latter was much worse than the former. BFGSF-1 performed about as well as sized BFGS-like methods, even though BFGSF-1 does not employ a sizing technique. However, this tendency cannot be observed between DFPF-1 and the sized DFP-like methods. The sized BFGS-like and the sized DFP-like methods performed well for all the problems. Sizing techniques take effect for the DFP-like methods better than for the BFGS-like methods. Actually, as shown in the results of BFGSF-3a, b and BFGSF-4a, b, the BFGS-like methods seem sensitive to choosing the sign of β_k'. In addition, it is interesting that the BFGS-like methods with (3.4) perform well whether sizing techniques are employed or not. On the other hand, all the methods based on the SZ update performed very well except SZF-0. Their behavior made little difference whether either sizing techniques or the switching to the Gauss-Newton method is employed or not. On the whole, the SZ method performed better than our methods did for the zero and small residual problems.

Tables 4 and 5 show the results for the PEAK problems, which are large residual problems. For those problems, we compared the BFGS-like update with the SZ update. Roughly speaking, sizing techniques did not have much effect on the performance in the SZ method as well as our methods.

8. Conclusion

This paper is concerned with the structured quasi-Newton methods for nonlinear least squares problems. We review our factorized quasi-Newton updates, the BFGS-like and the DFP-like updates. It seems that Sheng Songbai and Zou Zhihong introduced their update without investigating the existence of such a formula in detail. So we examine closely the existence and construct an updating formula which contains the SZ update, though we use the same least change secant update as Sheng Songbai et al. Further, we apply sizing techniques to the SZ method and propose new sizing factors.

Finally, the numerical comparison between our factorized quasi-Newton methods and the structured quasi-Newton methods based on the SZ update was made. On the whole, the SZ method performed better than our methods did. This fact seems to be caused by the difference between their conditions imposing on L_{k+1} and ours. We consider only the secant condition (3.2), while, they notice the combination of the Newton equation (1.3) and the approximation model of $r(x)$ given in (4.1), in addition to (3.2). Our numerical results suggest that consideration of an approximation model of $r(x)$ takes effect on computational performance. It is expected to exploit an efficient approximation model of $r(x)$.

Acknowledgment. The authors would like to thank Professor Y. Oyanagi, Tsukuba University, for sending us the PEAK problem, which is one of the large residual problems, for our numerical experiments.

References

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

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

[3] Ben-Israel, A. and Greville, T.N.E.: *Generalized Inverses — Theory and Applications.* Robert E. Krieger Publishing Company, Huntington, 1980.

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

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

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

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

[8] Nakagawa, T. and Oyanagi, Y.: Softwares for nonlinear least-squares methods. *Jyohoshori*, Information Processing Society of Japan, Vol.23, No.5 (1982), 442–450, in Japanese.

[9] Sheng Songbai and Zou Zhihong: *A new secant method for nonlinear least squares problems.* Technical Report NANOG-1988-03, Nanjing University, 1988.

[10] Yabe, H. and Takahashi, T.: Structured quasi-Newton methods for nonlinear least squares problems. *TRU Mathematics*, Vol.24, No.2 (1988), 195–209.

[11] Yabe, H. and Takahashi, T.: Factorized quasi-Newton methods for nonlinear least squares problems. *Mathematical Programming*, Vol.51, No.1 (1991), 75–100.

Table 1. Test Problems

Abbrebiated Name	Name of Test Problem	m	n	Starting Point	Residual
WATSON6	Watson Problem with 6 variables	31	6	(0, 0, ..., 0)	(S)
WATSON9	Watson Problem with 9 variables	31	9	(0, 0, ..., 0)	(S)
WATSON12	Watson Problem with 12 variables	31	12	(0, 0, ..., 0)	(S)
WATSON20	Watson Problem with 20 variables	31	20	(0, 0, ..., 0)	(S)
ROSENBROCK	Rosenbrock Problem	2	2	(-1.2, 1.0)	(Z)
HELIX	Helical Valley Problem	3	3	(-1, 0, 0)	(Z)
POWELL	Powell's Singular Problem	4	4	(3, -1, 0, 1)	(Z)
BEALE	Beale Problem	3	2	(0.1, 0.1)	(Z)
FRDSTEIN1	Freudenstein and Roth Problem	2	2	(6, 6)	(Z)
FRDSTEIN2	Freudenstein and Roth Problem	2	2	(15, -2)	(L)
BARD	Bard Problem	15	3	(1, 1, 1)	(S)
BOX	Box Problem	10	3	(0, 10, 20)	(Z)
KOWALIK	Kowalik Problem	11	4	(0.25 , 0.39, 0.415, 0.39)	(S)
OSBORNE1	Osborne Problem	33	5	(0.5, 1.5 , -1.0, 0.01, 0.02)	(S)
OSBORNE2	Osborne Problem	65	11	(1.3, 0.65, 0.65, 0.7, 0.6 , 3.0 , 5.0, 7.0 , 2.0 , 4.5 , 5.5)	(S)
JENNRICH	Jennrich Problem	10	2	(0.3, 0.4)	(L)
' PEAK	Peak Problem	51	5	(q, 2, 6, 3.5, 0.1) q = -2, -1, ..., 8	(L)

Table 2. Number of Iterations

	BFGS F-0	BFGS F-1	BFGS F-2a	BFGS F-2b	BFGS F-3a	BFGS F-3b	BFGS F-4a	BFGS F-4b
WATSON6	19	15	14	15	17	19	12	20
WATSON9	28	28	26	21	31	30	23	31
WATSON12	19	13	14	7	19	27	11	30
WATSON20	18	16	15	11	21	25	11	22
ROSENBROCK	29	22	13	16	16	13	15	12
HELIX	26	24	16	15	20	34	14	19
POWELL	20	14	14	14	14	14	14	14
BEALE	13	9	8	8	8	11	8	9
FRDSTEIN1	9	6	6	6	6	6	6	6
FRDSTEIN2	9	7	7	7	11	27	10	31
BARD	20	9	8	8	8	8	8	7
BOX	10	5	5	5	5	5	5	5
KOWALIK	14	10	9	9	8	12	9	10
OSBORNE1	43	27	18	18	16	17	15	26
OSBORNE2	22	24	15	15	16	16	17	14
JENNRICH	10	11	9	12	11	102*	10	16

Table 2. (Continued)

	DFP F-0	DFP F-1	DFP F-2a	DFP F-2b	DFP F-3a	DFP F-3b	DFP F-4a	DFP F-4b
WATSON6	76	27	8	8	9	9	9	8
WATSON9	126	47	23	23	21	21	21	21
WATSON12	153*	25	6	6	8	9	8	7
WATSON20	79	25	6	6	8	9	8	7
ROSENBROCK	500*	54	19	19	18	17	19	21
HELIX	283	34	12	12	12	14	11	15
POWELL	20	14	14	14	14	14	14	14
BEALE	19	11	8	8	7	12	7	11
FRDSTEIN1	8	6	6	6	6	6	6	6
FRDSTEIN2	10	6	6	6	8	10	8	13
BARD	17	9	8	8	8	8	8	8
BOX	15	5	5	5	5	5	5	5
KOWALIK	62	10	9	9	8	10	8	8
OSBORNE1	332*	57	139	139	21	19	20	26
OSBORNE2	40	41	13	13	13	15	13	15
JENNRICH	500*	89	9	9	8	8	8	8

Table 2. (Continued)

	SZ F-0	SZ F-1	SZ F-2a	SZ F-2b	SZ F-3a	SZ F-3b	SZ F-4a	SZ F-4b	SZ G-N1	SZ G-N2	SZ G-N3
WATSON6	9	10	7	7	7	7	7	7	8	10	10
WATSON9	20	20	19	19	19	20	20	19	19	20	20
WATSON12	6	6	5	5	5	5	5	5	5	6	6
WATSON20	6	6	5	5	5	5	5	5	5	6	6
ROSENBROCK	24	14	14	14	14	14	14	14	13	14	14
HELIX	7	11	11	11	11	11	11	11	10	11	11
POWELL	14	10	10	10	10	10	10	10	9	10	10
BEALE	12	9	7	7	8	8	9	9	7	9	9
FRDSTEIN1	7	6	6	6	6	6	6	6	6	6	6
FRDSTEIN2	99*	27	35[g]	35[g]	44[g]	93[g]	26[g]	19[g]	27	27	27
BARD	23	7	5	5	5	6	5	5	6	7	7
BOX	7	5	5	5	5	5	5	5	4	5	5
KOWALIK	13	10	8	8	9	10	9	9	19	10	10
OSBORNE1	42	33	28	28	18	19	18	17	6	33	33
OSBORNE2	24	21	14	14	12	14	12	13	21	21	21
JENNRICH	13	9	7	7	7	36	7	15	9	9	9

[g]: the global minimum is obtained.

Table 3. Number of Vector Valued Function Evaluations

	BFGS F-0	BFGS F-1	BFGS F-2a	BFGS F-2b	BFGS F-3a	BFGS F-3b	BFGS F-4a	BFGS F-4b
WATSON6	147	117	114	124	134	173	103	174
WATSON9	297	295	280	229	339	365	251	380
WATSON12	267	187	204	113	283	448	167	493
WATSON20	406	362	345	261	495	629	263	543
ROSENBROCK	96	78	48	88	79	59	77	57
HELIX	122	127	95	95	121	238	88	128
POWELL	105	75	75	75	75	75	75	75
BEALE	50	39	36	36	35	46	35	38
FRDSTEIN1	30	21	21	21	21	21	21	21
FRDSTEIN2	33	31	31	31	102	442	93	520
BARD	102	41	37	37	37	37	37	33
BOX	45	24	24	24	24	24	24	24
KOWALIK	94	61	56	56	51	74	56	61
OSBORNE1	274	181	128	128	118	127	107	219
OSBORNE2	286	310	200	200	216	222	232	194
JENNRICH	70	57	64	76	59	2013*	55	197

Table 3. (Continued)

	DFP F-0	DFP F-1	DFP F-2a	DFP F-2b	DFP F-3a	DFP F-3b	DFP F-4a	DFP F-4b
WATSON6	541	196	63	63	70	70	70	63
WATSON9	1279	480	240	240	220	220	220	220
WATSON12	2005*	342	91	91	117	130	117	104
WATSON20	1682	547	147	147	189	210	189	168
ROSENBROCK	1521*	180	68	68	73	69	80	82
HELIX	1142	143	60	60	59	66	56	74
POWELL	105	75	75	75	75	75	75	75
BEALE	65	42	33	33	30	45	30	41
FRDSTEIN1	27	21	21	21	21	21	21	21
FRDSTEIN2	33	21	21	21	39	49	38	63
BARD	77	41	37	37	37	37	37	37
BOX	64	24	24	24	24	24	24	24
KOWALIK	324	60	55	55	50	61	50	50
OSBORNE1	2003*	349	843	843	141	130	141	176
OSBORNE2	494	507	170	170	171	204	171	204
JENNRICH	1514*	272	32	32	29	29	29	29

Table 3. (Continued)

	SZ F-0	SZ F-1	SZ F-2a	SZ F-2b	SZ F-3a	SZ F-3b	SZ F-4a	SZ F-4b	SZ G-N1	SZ G-N2	SZ G-N3
WATSON6	70	77	56	56	56	56	56	56	63	77	77
WATSON9	210	210	200	200	200	210	210	200	200	210	210
WATSON12	91	91	78	78	78	78	78	78	78	91	91
WATSON20	147	147	126	126	126	126	126	126	126	147	147
ROSENBROCK	126	61	61	61	61	61	61	61	57	61	61
HELIX	34	51	51	51	51	51	51	51	47	51	51
POWELL	77	55	55	55	55	55	55	55	50	55	55
BEALE	44	35	30	30	32	32	35	36	29	135	35
FRDSTEIN1	34	22	22	22	22	22	22	22	22	22	22
FRDSTEIN2	2020*	468	583[g]	583[g]	748[g]	1800[g]	384[g]	251[g]	468	468	468
BARD	107	32	24	24	24	28	24	24	28	32	32
BOX	35	24	24	24	24	24	24	24	20	24	24
KOWALIK	75	59	49	49	55	60	55	55	103	59	59
OSBORNE1	262	210	199	199	148	139	137	128	44	210	210
OSBORNE2	312	272	187	187	163	186	163	173	272	272	272
JENNRICH	44	32	30	30	26	616	26	109	32	32	32

[g]: the global minimum is obtained.

Table 4. Number of Iterations for PEAK Problem

	q=-2	q=-1	q=0	q=1	q=2	q=3	q=4	q=5	q=6	q=7, 8
BFGS:F-0	21#	13	13#	9	6	6	10	12#	16#	—
F-1	14	21	13	8	4	5	9	16	21	—
F-2a	13	17	9	7	4	5	7	12	12	—
F-2b	13	17	9	7	4	5	7	13	12	—
F-3a	18	26	9	7	4	5	7	—	12	—
F-3b	13#	43	8	7	4	5	7	19	11	—
F-4a	14	19	10	7	4	5	7	—*	12	—
F-4b	14	13	9	7	5	5	7	113	13	—
S-Z:F-0	19	14	9	9	6	6	8	12	16	—
F-1	19	14	13	8	4	5	8	15	16	—
F-2a	14	9	9	6	4	5	7	12	11	—
F-2b	14	9	9	6	4	5	7	12	11	—
F-3a	16	9	9	7	4	5	7	9	13	—
F-3b	12	9	8	6	4	5	6	16	11	—
F-4a	14#	9	9	7	4	5	7	12	13	—
F-4b	16	9	9	6	5	5	7	14	14	—

: the negative Γ is obtained.
—: another stationary point is obtained.
(Note) For the cases of q=7 and 8, another stationary point is obtained by all the methods.

Table 5. Number of Vector Valued Function Evaluations for PEAK Problem

	q=-2	q=-1	q=0	q=1	q=2	q=3	q=4	q=5	q=6	q=7, 8
BFGS:F-0	156#	97	94#	62	42	42	70	89#	124#	—
F-1	106	150	91	54	30	36	62	123	170	—
F-2a	101	134	65	48	30	36	50	104	102	—
F-2b	101	134	65	48	30	36	50	110	102	—
F-3a	151	299	67	48	30	36	50	—	98	—
F-3b	103#	574	59	48	30	36	50	190	99	—
F-4a	115	205	73	48	30	36	50	—*	97	—
F-4b	116	115	66	48	36	36	51	1756	112	—
S-Z:F-0	127	92	60	60	42	42	54	78	108	—
F-1	127	94	86	55	30	36	55	100	111	—
F-2a	109	63	61	43	30	36	49	87	80	—
F-2b	109	63	61	43	30	36	49	89	80	—
F-3a	127	62	63	49	30	36	49	62	100	—
F-3b	87	62	55	43	30	36	43	141	81	—
F-4a	108#	62	63	49	30	36	49	86	100	—
F-4b	131	62	61	43	36	36	49	104	108	—

Hiroshi YABE:
Faculty of Engineering,
Science University of Tokyo,
1-3, Kagurazaka, Shinjuku-ku,
Tokyo, 162, Japan

Toshihiko TAKAHASHI:
Information Processing Center,
Kajima Corporation,
2-7, Motoakasaka 1-Chome,
Minato-ku, Tokyo, 107, Japan