NOVI SAD J. MATH.
VOL. 26, NO. 1, 1996, 115–127

QUASI-NEWTON'S METHOD WITH CORRECTION

Dragoslav Herceg, Nataša Krejić
Institute of Mathematics, Faculty of Science, University of Novi Sad
Trg Dositeja Obradovića 4, 21000 Novi Sad, Yugoslavia

Zorana Lužanin
Faculty of Engineering, University of Novi Sad
Trg Dositeja Obradovića 6, 21000 Novi Sad, Yugoslavia

Abstract

The most renowned method for solving nonlinear systems is the Newton's method. Its importance is primarily theoretical while practical application involves certain problems. The method proposed in this paper is aimed at eliminating one of those problems, i.e. the need to solve the linear (Newtonian) system in each iteration. This is made possible by retaining the same matrix of the Newtonian system while modifying only the right-side vector in each iteration. The proposed method can be considered as the improvement of the fixed Newton's method.

AMS Mathematics Subject Classification (1991): 65H10
Key words and phrases: nonlinear systems, Newton's method, Quasi-Newton's method

1. Introduction

We consider the problem of solving

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

where $F : \mathbb{R}^n \to \mathbb{R}^n$ is a nonlinear mapping of a special form

$$F(x) = Ax + G(x) \tag{2}$$

with linear part $A \in \mathbb{R}^{n,n}$ and nonlinear $G : \mathbb{R}^n \to \mathbb{R}^n$. A typical example is a system arising from the discretization of boundary value problem for ordinary or partial differential equations. In this case matrix A usually has some special properties like symmetry, diagonal dominance, good sparsity, see [9] for example, and solution of the linear system of equation with such a matrix is not a difficult task. The main idea in this paper is to exploit these properties in the application of Newton's method, i.e. we propose a quasi-Newton's method which generates iterations from the linear systems with the same matrix.

The most popular method for solving (1) is Newton's method. For a starting approximation $x^0 \in \mathbb{R}^n$ it generates an iterative sequence $\{x^k\}$ according to the rule

$$J\left(x^k\right) s^k = -F\left(x^k\right), \tag{3}$$

$$x^{k+1} = x^k + s^k, \quad k = 0, 1, \ldots, \tag{4}$$

where $J(x)$ is the Jacobian matrix of F. Newton's method is locally $q-$quadratically convergent but requires the calculation of $J\left(x^k\right)$ and solution of the linear system (3) at every iteration. Many modifications were introduced to overcome these difficulties. Basically, we can divide those modifications into two wide classes: inexact Newton's methods, which solve the linear Newtonian system only approximately, see [3], and quasi-Newton's methods where (3) is replaced with quasi-Newtonian linear system. Both modifications try to save computational work in individual iteration and maintain some convergence properties of Newton's method. In most quasi-Newton's method an approximation for $J\left(x^k\right)$ or $J\left(x^k\right)^{-1}$ is calculated such that computational effort is reduced. On the other hand these methods have $q-$linear or $q-$superlinear local convergence, see [4], [7], [8].

In this paper we introduce a new quasi-Newton's method based on the idea to use good properties of the matrix A from (2). We prove local $q-$linear convergence and report numerical results which show its good behaviour.

2. Description of the new method

For an approximation x^k the next iteration of Newton's method for solving (1), where $F(x) = Ax + G(x)$, is given by

$$\begin{aligned} \left(A + G'\left(x^k\right)\right) s^k &= -F\left(x^k\right), \\ x^{k+1} &= x^k + s^k, \end{aligned}$$

where $G'(x)$ is the Jacobian matrix for the nonlinear mapping $G(x)$.

Let m be a positive integer, $x^0 \in \mathbb{R}^n$ initial approximation and $H_0 \in \mathbb{R}^{n,n}$. We define an iterative sequence by

(5) $$x^{k+1} = x^k - H_k F\left(x^k\right), \quad k = 0, 1, \dots$$

in the following way. If $k+1 \equiv 0 \pmod m$, we choose $H_{k+1} \in \mathbb{R}^{n,n}$ and the iteration is finished. If $k+1$ is not a multiply of m, then we solve

(6) $$As^k = -\left(E + G'\left(x^k\right)\right) F\left(x^k\right)$$

and define

(7) $$x^{k+1} = x^k + s^k.$$

So for $k+1 \equiv q \pmod m$, $q \geq 1$, we have

$$H_{k+1} = A^{-1}\left(E + G'\left(x^k\right)\right).$$

In this way we obtain the method which requires the solutions of m systems of linear equations with the same matrix. Contrary to the so-called fixed Newton's method, see [9], given with

$$J\left(x^0\right) s^k = -F\left(x^k\right), \quad x^{k+1} = x^k + s^k,$$

and the method of direct iterations, see [1],

$$Ax^{k+1} = -G\left(x^k\right),$$

we add a correction $-G'\left(x^k\right) F\left(x^k\right)$ in every iteration (for $k+1 \equiv q \pmod m$, $q \geq 1$). Practical experience suggests that in many cases correction $-\alpha G'\left(x^k\right) F\left(x^k\right)$, as we are going to see in Section 4, but it has no influence on the proof of convergence theorem, so we are going to consider the case $\alpha = 1$.

3. Convergence result

From now on, we denote by $\|\cdot\|$ 2-norm for vectors and matrices.

Assume that $F : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$, $F \in C^1(D)$, D is an open and convex set, $x^* \in D$ such that $F(x^*) = 0$ and

$$\|J(x) - J(x^*)\| \leq L\|x - x^*\|^p, \quad L, p > 0 \tag{8}$$

for all $x \in D$. We also suppose that $J(x^*)$ and A are nonsingular matrices,

$$\|J(x^*)^{-1}\| \leq M$$

and

$$\|A^{-1}\| \leq C.$$

Assumption (8) implies

$$\|F(x) - F(x^*) - J(x^*)(x - x^*)\| \leq L\|x - x^*\|^{p+1},$$

for all $x \in D$, see [9].

We are going to use the following lemma.

Lemma 1. *[6] (Banach Lemma) If $M \in \mathbb{R}^{n,n}$ with $\|M\| < 1$ then $E - M$ is nonsingular and*

$$\|(E - M)^{-1}\| \leq \frac{1}{1 - \|M\|}.$$

Theorem 2. *Let $r \in (0,1)$. There exists $\varepsilon = \varepsilon(r)$, $\delta = \delta(r)$ such that if $\|x^* - x^0\| \leq \varepsilon$ and $\left\|H_k - J(x^*)^{-1}\right\| \leq \delta$ whenever $k \equiv 0 \pmod{m}$ then the sequence $\{x^k\}$ is generated by (5)-(7) are well defined, converges to x^* and*

$$\left\|x^{k+1} - x^*\right\| \leq r\left\|x^k - x^*\right\|$$

for all $k = 0, 1, \ldots$.

Proof. Define $K = CL(1 + r)$ and for $\varepsilon, \delta > 0$ consider functions

$$\begin{aligned} b_0(\varepsilon, \delta) &= \delta, \\ b_i(\varepsilon, \delta) &= b_{i-1}(\varepsilon, \delta) + K\varepsilon^p, \quad i = 1, 2, \ldots, m - 1. \end{aligned}$$

Clearly, for all $\varepsilon, \delta > 0$ follows

$$0 < b_0(\varepsilon,\delta) < b_1(\varepsilon,\delta) < \ldots < b_{m-1}(\varepsilon,\delta)$$

and

$$\lim_{\varepsilon,\delta\to 0} b_i(\varepsilon,\delta) = 0, \quad i = 0,1,\ldots,m-1.$$

We can choose $\varepsilon = \varepsilon(r)$ and $\delta = \delta(r)$ such that

$$b_i(\varepsilon,\delta) + L\varepsilon^p < \frac{r}{M_1}, \quad i = 0,1,\ldots,m-1,$$

with $M_1 = \max\{\|J(x^*)\|, 2M\}$.

Assume that $\|x^0 - x^*\| \leq \varepsilon$ and $\left\|H_k - J(x^*)^{-1}\right\| \leq \delta$ whenever $k \equiv 0 \pmod m$. We are going to prove that if $k \equiv q \pmod m$ then H_k is nonsingular,

$$\left\|x^{k+1} - x^*\right\| \leq r\left\|x^k - x^*\right\|, \tag{9}$$

$$\left\|H_k - J(x^*)^{-1}\right\| \leq b_q(\varepsilon,\delta) \tag{10}$$

and

$$\|H_k\| \leq 2M \tag{11}$$

for all $q = 0,1,\ldots,m-1$.

For $k = 0$, by hypothesis,

$$\left\|H_0 - J(x^*)^{-1}\right\| \leq \delta = b_0(\varepsilon,\delta)$$

and

$$\begin{aligned}\|H_0\| &\leq \left\|J(x^*)^{-1}\right\| + \left\|H_0 - J(x^*)^{-1}\right\| \\ &\leq M + \delta \\ &\leq 2M.\end{aligned}$$

Moreover,

$$\begin{aligned}\left\|x^1 - x^*\right\| &= \left\|x^0 - x^* - H_0 F(x^0)\right\| \\ &= \left\|x^0 - x^* - H_0\left[F(x^0) - F(x^*) - J(x^*)(x^0 - x^*)\right]\right. \\ &\quad \left. - H_0 J(x^*)(x^0 - x^*)\right\|\end{aligned}$$

$$
\begin{aligned}
&\leq \left\|x^0 - x^* - H_0 J(x^*)(x^0 - x^*)\right\| \\
&\quad + \|H_0\| \left\|F(x^0) - F(x^*) - J(x^*)(x^0 - x^*)\right\| \\
&\leq \|E - H_0 J(x^*)\| \left\|x^0 - x^*\right\| + \|H_0\| L \left\|x^0 - x^*\right\|^{p+1} \\
&\leq \left(\|J(x^*)\| \left\|J(x^*)^{-1} - H_0\right\| + \|H_0\| L \left\|x^0 - x^*\right\|^p\right) \left\|x^0 - x^*\right\| \\
&\leq M_1 \left(b_0(\varepsilon, \delta) + L\varepsilon^p\right) \left\|x^0 - x^*\right\| \\
&\leq r \left\|x^0 - x^*\right\|.
\end{aligned}
$$

Consider now $k > 0$, $k \equiv q \,(\mathrm{mod}\, m)$. If $q = 0$ the proof of (9)-(11) is analogous to the case $k = 0$, so assume $q > 0$. Then

$$
\begin{aligned}
H_k &= A^{-1}\left(E + G'\left(x^k\right)\right) \pm A^{-1} G'\left(x^{k-1}\right) \\
&= A^{-1}\left(E + G'\left(x^{k-1}\right)\right) + A^{-1}\left(G'\left(x^k\right) - G'\left(x^{k-1}\right)\right),
\end{aligned}
$$

so

$$
\begin{aligned}
\left\|H_k - J(x^*)^{-1}\right\| &\leq \left\|H_{k-1} - J(x^*)^{-1}\right\| + \left\|A^{-1}\left(G'\left(x^k\right) - G'\left(x^{k-1}\right)\right)\right\| \\
&\leq b_{q-1}(\varepsilon, \delta) + \left\|A^{-1}\right\| \left(\left\|G'\left(x^k\right) - G'(x^*)\right\|\right. \\
&\quad \left. + \left\|G'(x^*) - G'\left(x^{k-1}\right)\right\|\right) \\
&\leq b_{q-1}(\varepsilon, \delta) + CL\left(\left\|x^k - x^*\right\|^p + \left\|x^{k-1} - x^*\right\|^p\right) \\
&\leq b_{q-1}(\varepsilon, \delta) + CL(1 + r)\varepsilon^p.
\end{aligned}
$$

Therefore

$$
\left\|H_k - J(x^*)^{-1}\right\| \leq b_q(\varepsilon, \delta) \leq \frac{r}{M_1} \leq \frac{1}{2M}
$$

and, by Lemma 1, H_k is nonsingular. Hence

$$
\begin{aligned}
\|H_k\| &\leq \left\|J(x^*)^{-1}\right\| + \left\|H_k - J(x^*)^{-1}\right\| \\
&\leq \left\|J(x^*)^{-1}\right\| + \frac{r}{M_1} \\
&\leq \left\|J(x^*)^{-1}\right\| + \|J(x^*)\|^{-1} \leq 2M.
\end{aligned}
$$

Finally,

$$\begin{aligned}
\left\|x^{k+1}-x^*\right\| &\le \left\|x^k-x^*-H_kF\left(x^k\right)\right\| \\
&= \left\|x^k-x^*-H_k\left[F\left(x^k\right)-F\left(x^*\right)-J\left(x^*\right)\left(x^k-x^*\right)\right]\right. \\
&\quad \left. -H_kJ\left(x^*\right)\left(x^k-x^*\right)\right\| \\
&\le \left\|E-H_kJ\left(x^*\right)\right\|\left\|x^k-x^*\right\| \\
&\quad +\left\|H_k\right\|\left\|F\left(x^k\right)-F\left(x^*\right)-J\left(x^*\right)\left(x^k-x^*\right)\right\| \\
&\le M_1\left(b_q\left(\varepsilon,\delta\right)+L\varepsilon^p\right)\left\|x^k-x^*\right\| \\
&\le r\left\|x^k-x^*\right\|.
\end{aligned}$$

Therefore, (9) is also proved, so $\left\{x^k\right\}$ is a convergent sequence and

$$\lim_{k\to\infty} x^k = x^*.$$

4. Practical implementation and numerical results

The first example we present here is a nonlinear system arising from the standard five point discretization of a Poisson equation

$$-\Delta u+\phi\left(u\right)=f\left(x,y,u\right)$$

in the square $[0,1]\times[0,1]$ with the boundary condition $u\left(x,y\right)=0$. We report results for $\phi\left(u\right)=u^3$. This example is the representative of a very important class of nonlinear systems which come from the discretization of PDEs arising in many mathematical models. The number of divisions of the interval $[0,1]$ on both axis is denoted by N, so the number of variables is $n=\left(N-1\right)^2$. Linear part of the equation gives the matrix A which is a block tridiagonal matrix, and in each step of the Newton's method $J\left(x^k\right)$ is also a block tridiagonal matrix. In the implementation of the proposed new method and Newton's method we stored LU factorization of A and $J\left(x^k\right)$ (see [5]). The stopping criterion was

$$\left\|F\left(x^k\right)\right\|\le 10^{-5}$$

and the initial point was $x^0 = [0,0,\dots,0]^{\mathrm{T}}$. The results are reported in Table 1, each experiment is represent by a pair (k,t) where k is the number of iterations and t is the computer CPU time in seconds and NM denotes Newton's method and CM denotes the method proposed in this paper. We can see that CM requires a slightly larger number of iterations than the Newton's method, but at the same time takes significantly less CPU time.

n	NM	CM
49	$(8, 34.7)$	$(14, 39.2)$
225	$(8, 156.0)$	$(9, 101.2)$
961	$(7, 851.7)$	$(8, 399.6)$
3969	$(7, 6589.3)$	$(7, 1845.4)$

Table 1.

As we mentioned in Section 2, practical experience suggests that in many cases correction should be $-\alpha G'\left(x^k\right)F\left(x^k\right)$, i.e. the iterations become

$$\begin{aligned} As^{k+1} &= -\left(E + \alpha G'\left(x^k\right)\right)F\left(x^k\right), \\ x^{k+1} &= x^k + s^k. \end{aligned}$$

We are going to demonstrate the influence of real parameter α in the following simple example given in Brown, [2].

$$\begin{aligned} f_1(x) &= -1 + \prod_{k=1}^{n} x_k, \\ f_i(x) &= -(n+1) + x_i + \sum_{k=1}^{n} x_k, \quad i = 2,3,\dots,n. \end{aligned}$$

In this example the choice of a nonsingular matrix A that we need for the proposed method is not naturally determined as it is in the case of PDEs, because the linear part of the considered mapping gives matrix with all zeros in first row,so it has to be modified. For $n = 4$ we choose

$$A_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix} \quad \text{and} \quad A_2 = J\left(x^0\right).$$

Obviously, with $A_2 = J\left(x^0\right)$ and $\alpha = 0$ we obtain fixed Newton's method and with A_1 the method of direct iterations, but A_2 is not an M-matrix and convergence conditions given in [1] are not satisfied. Newton's method required 8 iterations with the termination criterion

$$\left\|F\left(x^k\right)\right\| < 10^{-8}$$

and starting approximation $x^0 = [0.9, 0.9, 0.9, 0.9]^\top$ and took 0.8 seconds CPU time. The best result obtained with the correction method for the same starting approximation and termination criterion was 5 iterations which required 0.33 seconds CPU time. The results are reported in Table 2, with D denoting divergence.

α	1	0.4	0.3	0.2	0.1	0	-0.1	-0.2	-0.2	-0.4	1
A_1	D	D	D	D	12	7	5	7	8	D	D
A_2	53	28	22	18	14	15	14	12	10	7	10

Table 2.

Influence of the parameter α can be better explained if we consider the correction method as a sort of inexact Newton's method

$$\begin{aligned} J\left(x^k\right) s^k &= -F\left(x^k\right) + r^k, \\ x^{k+1} &= x^k + s^k, \\ \frac{\|r^k\|}{\|F(x^k)\|} &\le \eta_k < \eta < 1, \end{aligned}$$

see [3]. In our case

$$r^k = G'\left(x^k\right)\left(\alpha E + A^{-1} + \alpha A^{-1} G'\left(x^k\right) F\left(x^k\right)\right)$$

and consequently if $\eta_k \le 1$ with

$$\eta_k = \frac{\left\|G'\left(x^k\right)\left(\alpha E + A^{-1} + \alpha A^{-1} G'\left(x^k\right) F\left(x^k\right)\right)\right\|}{\|F(x^k)\|} \tag{12}$$

convergence follows from the theorem from [3]. From (12) we see that η_k is a polynomial in α, so if we determine α as

$$\min_{\alpha}\left\|G'\left(x^k\right)\left(\alpha E + A^{-1} + \alpha A^{-1} G'\left(x^k\right) F\left(x^k\right)\right)\right\|, \tag{13}$$

we can expect fast convergence. Numerical results confirm this prediction. Obviously, for evaluating such α we need A^{-1} explicitly, which is a problem for systems with larger dimension, but in these cases we determine α as

$$\min_{\alpha} \left\| E + \alpha A + \alpha G'\left(x^k\right) F(x^k) \right\| \tag{14}$$

In our example A^{-1} is calculated exactly and results with optimal α are presented in Table 3. As η_k is very close to 0, we practically obtain Newton's iterative sequence. Of course, calculation of optimal α also requires some computational effort, but less than the application of Newton's method. In Table 3 $k, (A_i, \alpha_k)$, η_k denote sequential of iteration, optimal α_k for A_i, $i = 1, 2$ and η_k from (12) respectively. Termination criterium was satisfied after 10 iterations for the method with A_1 and 8 iterations for the method with A_2.

k	(A_1, α_k)	η_k	(A_2, α_k)	η_k
1	1.	0.490	1.	0.
2	-0.34	$5.06 \cdot 10^{-15}$	-94.87	$2.37 \cdot 10^{-16}$
3	-0.69	$2.13 \cdot 10^{-15}$	3.57	0.
4	-1.35	$4.7 \cdot 10^{-15}$	10.5	0.
5	-2.55	$4.47 \cdot 10^{-15}$	21.48	$1.14 \cdot 10^{-15}$
6	-4.29	$1.27 \cdot 10^{-16}$	25.58	0.
7	-5.66	0.	25.73	$4.09 \cdot 10^{-16}$
8	-5.98	$2.68 \cdot 10^{-19}$	33.91	$8.16 \cdot 10^{-16}$
9	-5.99	0.		
10	-6.0	0.		

Table 3.

All reported results were obtained using *Mathematica* on the PC platform (486-DX120,16MB RAM).

The example from Table 2 shows that the proposed method improves the fixed Newton method ($\alpha = 0,\ A = A_2$) with respect to the number of iterations. We will show on another example that the proposed method also expands the area for the selection of initial iteration which is very important when solving nonlinear systems by use of iterative methods.

We consider a problem of solving this system of nonlinear equations:

$$f_1(x_1,x_2) = x_1^3 - 3x_1x_2^2 - 1;$$
$$f_2(x_1,x_2) = 3x_1^2x_2 - x_2^3;$$

which has three roots, $x_1^* = [1,0]^T$, $x_2^* = [0.5, \sqrt{3}\,0.5]^T$ i $x_3^* = [0.5, -\sqrt{3}\,0.5]^T$. For the fixed Newton's method and the new method, a test for convergence was done where the initial approximation x^0 was selected frot the recangle $D = [-3.5, 3.5] \times [-2.5, 2.5]$.

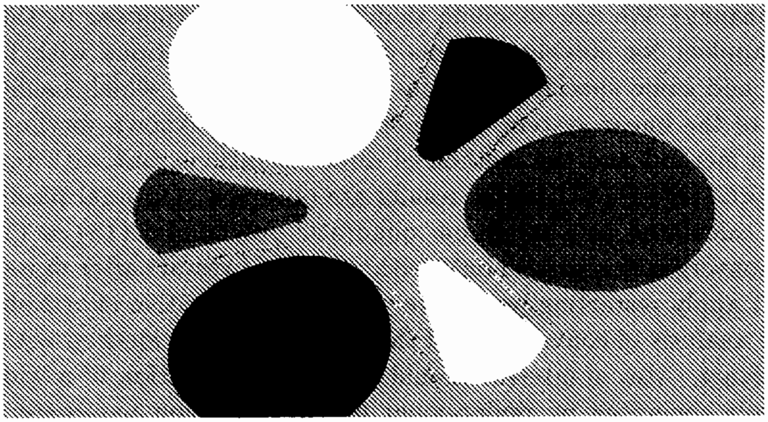


Figure 1: Fixed Newton's method

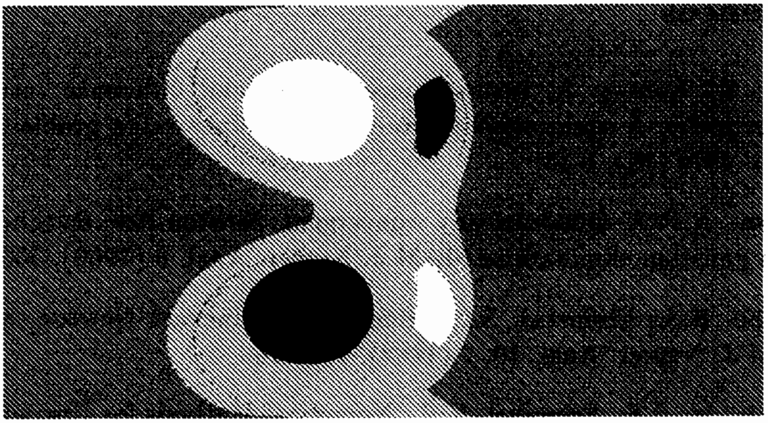
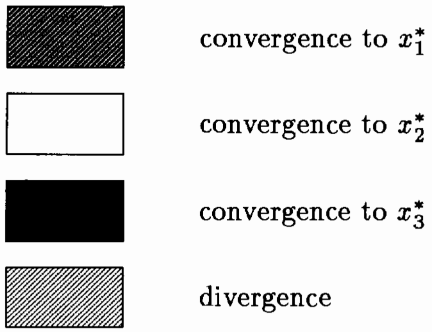


Figure 2: New method with $\alpha = -0.1$ and $m = \infty$



Summary: We can conclude that the method proposed in this paper can be applied for solving systems of nonlinear equations with large dimension and some special properties of its linear part which facilitate the solution of the linear system. It generalizes the well-known methods, fixed Newton's method and the method of direct iterations, but with correction vector $-\alpha G'\left(x^k\right) F\left(x^k\right)$ better results are obtained. Optimal parameter α is easily calculated as a minimum quadratic polynomial (13) or (14). Although convergence is proved only with periodical restart after m iterations, we didn't apply restarting procedure. Future investigative efforts should be aimed at analysis of convergence of the newly proposed method without the use of restart. Also of interest should be the consideration of a wider class of methods generated by replacement of the relaxation parameter α with the diagonal relaxation matrix $L = \text{diag}(\alpha_1, \cdots, \alpha_n)$.

References

[1] Bohl, E., Lorenz, J., Inverse monotonicity and difference schemes of higher order. A summary for two-point boundary value problems, Aeq. Math. 1979 (19), 1-13.

[2] Brown, R.P.,A Quadratically convergent Newton-like method based upon gaussian eliminations,SIAM J. Numer. Anal. 6 (1969), 560-569.

[3] Dembo, R.S., Eisenstat, S.C., Steihaug, T., Inexact Newton methods, SIAM J. Numer. Anal. 19, 2 (1982), 400-408.

[4] Dennis, Jr., J.E., Schnabel, R.B., Numerical Methods for Unconstrained Optimization and Nonlinear Equations, Prentice Hall, Englewood Cliffs, N.Y., 1983.

[5] Golub, G. H., Van Loan, Ch. F., Matrix Computations. The John Hopkins University Press, Baltimore, 1983.

[6] Kelley, C.T., Iterative Methods for Linear and Nonlinear Equations, Frontiers in Applied Mathematics 16, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1995.

[7] Martínez, J.M., A Quasi-Newton method with modification of one column per iteration,Computing 3 (1984), 353-362.

[8] Martínez, J.M.,A Family of quasi-Newton methods for nonlinear equations with direct secant updates of matrix factorizations,SIAM J. Numer. Anal. 27(1990), 1034-1049.

[9] Ortega, J.M., Rheinboldt, W.C., Iterative Solution of Nonlinear Equations of Several Variables, Academic Press, New York, 1970.

Received by the editors June 1, 1996