Computing 33, 353 – 362 (1984)

Computing

A Quasi-Newton Method with Modification of One Column per Iteration

J. M. Martínez, Campinas

Received April 2, 1984

Abstract — Zusammenfassung

A Quasi-Newton Method with Modification of One Column per Iteration. In this paper we introduce a new Quasi-Newton method for solving nonlinear simultaneous equations. At each iteration only one column of B_k is changed to obtain B_{k+1}. This permits to use the well-known techniques of Linear Programming for modifying the factorization of B_k. We present a local convergence theorem for a restarted version of the method. The new algorithm is compared numerically with some other methods which were introduced for solving the same kind of problems.

AMS Subject Classification: 65H10.

Key words: Nonlinear systems, Quasi-Newton methods.

Ein Quasi-Newton-Verfahren mit Veränderung einer Spalte pro Iteration. Wir stellen ein neues Quasi-Newton-Verfahren vor zur Lösung von nichtlinearen simultanen Gleichungen. Bei jeder Iteration wird lediglich eine Spalte von B_k verändert, um B_{k+l} zu erhalten. Dies erlaubt, wohlbekannte Techniken der Linearen Programmierung zur Faktorisierung von B_k zu benützen. Wir beweisen einen Satz über die lokale Konvergenz für die Methode. Der neue Algorithmus wird mit anderen bezüglich seiner numerischen Eigenschaften verglichen.

1. Introduction

Many problems require the numerical solution of a system of n nonlinear equations in n unknowns:

$$\text{given } F : \mathbb{R}^n \to \mathbb{R}^n, \text{ find } x^* \in \mathbb{R}^n \text{ such that } F(x^*) = 0. \tag{1}$$

The numerical solution of this problem is usually iterative, proceeding at each iteration from an estimate x^k of x^* to a better estimate x^{k+1}. The Newton step s_N^k is the solution of the linear system

$$F'(x^k)\, s_N^k = -F(x^k),$$

where

$$(F'(x))_{ij} = \frac{\partial f_i}{\partial x_j}(x),$$

f_i the i-th component function of F, x_j the j-th component of the vector x, $(F'(x))_{ij}$ the component of the matrix $F'(x)$ in row i and column j.

When analytic derivatives are not available, secant methods ([11, 13, 16, 21, 23]) and Quasi-Newton ($Q-N$) methods ([3, 4, 7]) represent a significant improvement over the discretized version of the classical Newton's method. $Q-N$ methods are based in the formulae:

$$\begin{aligned} B_k s^k &= -F(x^k), \\ x^{k+1} &= x^k + s^k, \\ B_{k+1} s^k &= y^k, \end{aligned} \tag{2}$$

where $y^k = F(x^{k+1}) - F(x^k)$, for all $k = 0, 1, 2, \ldots$.

(2) is called the Fundamental Equation of $Q-N$ Methods. If $n \geq 2$ and $s^k \neq 0$, many matrices will obey (2). If the Jacobian $F'(x)$ has special properties, such as symmetry of sparsity, we may restrict the choice of B_{k+1} to the set of matrices which have these properties. Schubert ([22]), Dennis-Marwil ([5]) and Martínez [15] introduced methods which allow to keep the sparsity structure of the Jacobian matrix and therefore, are able to deal with large-scale problems.

In this paper we introduce a new Quasi-Newton method where at each iteration only one column of B_k is changed to obtain B_{k+1}. Provided we store a suitable factorization of B_k, the factorization of B_{k+1} is obtained using the classical procedures used in Linear Programming ([1, 10]) even in the sparse case. The sparsity structure of $F'(x)$ remains and the equation (2) is satisfied in most cases.

In Section 2 we define the method and prove a local convergence theorem. In Section 3 we describe the computational implementation and present some numerical experiences. Finally, in Section 4 we state some conclusions and suggest lines for future research.

2. Local Convergence Theorem

General Hypotheses: Let $F: D \subset \mathbb{R}^n \to \mathbb{R}^n$, $F \in C^1(D)$, D an open and convex set. Let $F(x^*) = 0$ and $F'(x^*)$ be a nonsingular $n \times n$ matrix. We assume that for all $x \in D$

$$\| F'(x) - F'(x^*) \| \leq K \, \| x - x^* \|^p, \quad K, p > 0.$$

$\| \cdot \|$ will denote the 2-norm throughout the paper.

Lemma 2.1: *For all $u, v \in D$*

$$\begin{aligned} &\| F(v) - F(u) - F'(x^*)(v-u) \| \leq \\ &\leq K \max \{ \| v - x^* \|^p, \ \| u - x^* \|^p \} \ \| v - u \|. \end{aligned} \tag{3}$$

Proof: See [6].

The Proposed Method: Let $x^0 \in D$ an arbitrary initial point, $\alpha \in (0, 1/\sqrt{n})$, m a positive integer, B_0 a nonsingular $n \times n$ matrix. We define recursively a sequence of points (x^k) and matrices B_k by

$$\begin{aligned} x^{k+1} &= x^k - B_k^{-1} F(x^k) \\ s^k &= x^{k+1} - x^k \\ y^k &= F(x^{k+1}) - F(x^k). \end{aligned}$$

If $k+1$ is not a multiple of m, let j be such that $|s_j^k| > \alpha \| s^k \|$. Then define $B_{k+1} = (b_{il}^{k+1})$ as being equal to $B_k = (b_{il}^k)$ except perhaps at column j. Set I_j^k be a subset of $\{1, \ldots, n\}$. Then if $i \in I_j^k$, we define

$$b_{ij}^{k+1} = (y_i^k - \sum_{l \neq j} b_{il}^k s_l^k)/s_j^k$$

and

$$b_{ij}^{k+1} = b_{ij}^k \text{ if } i \notin I_j^k.$$

It is easy to see that if $k+1$ is not a multiple of m and $I_j^k = \{1, \ldots, n\}$ then the equation (2) is verified. The choice of I_j^k is made in order to preserve the structure of B_k. Notice that from $\| s^k \| / |s_j^k| < 1/\alpha$ it follows that $|s_i^k| / |s_j^k| < 1/\alpha$ for all $i = 1, \ldots, n$. We refer to the choice of B_{k+1} when $k+1 \equiv 0 \mod (m)$ in the following theorem.

Theorem 2.1: *Let $r \in (0,1)$. Then there exist $\varepsilon = \varepsilon(r)$ and $\delta = \delta(r)$ such that if $\| x^0 - x^* \| \leq \varepsilon$ and $\| B_k - F'(x^*) \| \leq \delta$ whenever $k \equiv 0 \mod (m)$ then the sequences (x^k) and (B_k) are well defined and if $F(x^k) \neq 0$ then $\| x^{k+1} - x^* \| \leq r \| x^k - x^* \|$ for all $k = 0, 1, 2, \ldots$.*

Proof: Define $c_1 = \sqrt{n}\, K/\alpha$, $c_2 = n^{3/2}/\alpha$. Given $\varepsilon, \delta > 0$, define $b_0, b_1, \ldots, b_{m-1} = b(\varepsilon, \delta)$ by $b_0 = \delta$, $b_{k+1} = c_2 b_k + c_1 \varepsilon^p$, $k = 0, 1, \ldots, m-2$. We verify that $\lim_{\varepsilon, \delta \to 0} b(\varepsilon, \delta) = 0$ and $0 < b_0 < b_1 < \ldots < b_{m-1}$.

Let ε and δ be such that

$$b(\varepsilon, \delta) + K \varepsilon^p \leq r/(2 \| F'(x^*)^{-1} \|) \tag{4}$$

We shall prove by induction on k that if $k \equiv q \pmod m$ then B_k is nonsingular (then x^{k+1} is well-defined),

$$\| x^{k+1} - x^* \| \leq r \| x^k - x^* \|, \quad \| B_k - F'(x^*) \| \leq b_q$$

and

$$\| B_k^{-1} \| \leq 2 \| F'(x^*)^{-1} \|.$$

For $k = 0$, by hypotheses,

$$\| B_0 - F'(x^*) \| \leq \delta = b_0.$$

Now, by (4),

$$\delta + K \varepsilon^p \leq r/(2 \| F'(x^*)^{-1} \|).$$

Then

$$\delta \leq 1/(2 \| F'(x^*)^{-1} \|).$$

Therefore, by Banach lemma of perturbation (see [19])

$$\| B_0^{-1} \| \leq 2 \| F'(x^*)^{-1} \|. \tag{5}$$

Then, the fact that $\| x^1 - x^* \| \leq r \| x^0 - x^* \|$ follows from (3) using classical arguments (see [4, 6]).

Consider now an arbitrary k. If $q = 0$ the proof is the same as in the case $k = 0$. If $q > 0$ let us prove first that $\| B_k - F'(x^*) \| \leq b_q$. We use $\| B_{k-1} - F'(x^*) \| \leq b_{q-1}$. Let us call $s = x^k - x^{k-1}$, $y = F(x^k) - F(x^{k-1})$. If $F(x^{k-1}) \neq 0$ we may suppose $s \neq 0$.

Put $B_k = (b_{il}^k)$, $B_{k-1} = (b_{il}^{k-1})$, $F'(x^*) = (b_{il}^*)$. Let j be the index of the column which is changed from B_{k-1} to B_k. Then $|s_j| > \alpha \| s \| > 0$.

So, if $i \in I_j^k$

$$b_{ij}^k = (y_i - \sum_{l \neq j} b_{il}^{k-1} s_l)/s_j =$$

$$= (y_i - \sum_{l \neq j} b_{il}^* s_l + \sum_{l \neq j} b_{il}^* s_l - \sum_{l \neq j} b_{il}^{k-1} s_l)/s_j.$$

Thus,

$$| b_{ij}^k - b_{ij}^* | \leq | y_i - \sum_{l=1}^{n} b_{il}^* s_l | / | s_j |$$

$$+ \sum_{l \neq j} | b_{il}^* - b_{il}^{k-1} | \, | s_l | / | s_j | \leq$$

$$\leq \| y - F'(x^*) s \| / | s_j | + \left(\sum_{l=1}^{n} | b_{il}^* - b_{il}^{k-1} | \right) / \alpha \leq$$

$$\leq \| y - F'(x^*) s \| / (\alpha \| s \|) + (n/\alpha) \| B_{k-1} - F'(x^*) \| \leq$$

$$\leq (K/\alpha) \| x^{k-1} - x^* \|^p + n \, b_{q-1}/\alpha \leq$$

$$\leq K \varepsilon^p / \alpha + n \, b_{q-1}/\alpha.$$

Of course, the same inequality holds for $i \notin I_j^k$, therefore,

$$\| B_k - F'(x^*) \| \leq \sqrt{n} \, K \varepsilon^p / \alpha + n^{3/2} b_{q-1}/\alpha$$
$$= c_1 \varepsilon^p + c_2 b_{q-1} = b_q. \tag{6}$$

Let us now prove that B_k^{-1} exists and $\| B_k^{-1} \| \leq 2 \| F'(x^*)^{-1} \|$. In fact, by (4) and (6)

$$\| B_k - F'(x^*) \| \leq b_q \leq b(\varepsilon, \delta) \leq b(\varepsilon, \delta) + K \varepsilon^p \leq$$

$$\leq r/(2 \| F'(x^*)^{-1} \|) < 1/(2 \| F'(x^*)^{-1} \|).$$

Then, by the perturbation lemma, B_k^{-1} exists and

$$\| B_k^{-1} \| \leq 2 \| F'(x^*)^{-1} \|.$$

Finally, by (3)–(6)

$$\| x^{k+1} - x^* \| = \| x^k - x^* - B_k^{-1} F(x^k) \| = \| x^k - x^* -$$

$$- B_k^{-1} [F(x^k) - F'(x^*)(x^k - x^*)] - B_k^{-1} F'(x^*)(x^k - x^*) \| \leq$$

$$\leq \| x^k - x^* - B_k^{-1} F'(x^*)(x^k - x^*) \| + \| B_k^{-1} \| \; \| F(x^k) - F(x^*)$$

$$- F'(x^*)(x^k - x^*) \| \leq \| I - B_k^{-1} F'(x^*) \| \; \| x^k - x^* \| +$$

$$+ \| B_k^{-1} \| \; K \| x^k - x^* \|^{p+1} \leq$$

$$\leq \| B_k^{-1} \| \; \| B_k - F'(x^*) \| + K \varepsilon^p \| B_k^{-1} \| \; \| x^k - x^* \| \leq$$

$$\leq 2 \| F'(x^*)^{-1} \| (b_p + K \varepsilon^p) \| x^k - x^* \| \leq$$

$$\leq 2 \| F'(x^*)^{-1} \| \left(b(\varepsilon, \delta) + K \varepsilon^p \right) \| x^k - x^* \| \leq$$

$$\leq r \| x^k - x^* \|.$$

□

3. Numerical Experiences

We implemented the method defined in the previous section using the $Q-R$ factorization of B_k. The modification of the $Q-R$ factorization was performed using the procedures described in [9, 16]. The implementation for large sparse problems uses the $L-U$ factorization and the classical updating techniques of Linear Programming ([1, 10]).

We claim that, if the new method is reliable for large sparse problems, it should also be reliable for small dense problems. We decided then to test our method for a number of classical small problems. Of course, the advantage of this approach is that there exist many well-known small problems in the literature with different kinds of difficulties (see [14, 16, 17]).

The test functions were the following:

Problem 1, $n=2$ (*Rosenbrock*)

$f_1(x)=10(x_2-x_1^2)$
$f_2(x)=1-x_1$
I: $x^0=(-1.2, 1)$
II: $10x^0$
III: $100x^0$

Problem 2, $n=2$ (*Freudenstein-Roth*)

$f_1(x)=-13+x_1+[(5-x_2)x_2-2]x_2$
$f_2(x)=-29+x_1+[(x_2+1)x_2-14]x_2$
I: $x^0=(0.5, -2)$
II: $10x^0$
III: $100x^0$

Problem 3, $n=5$ (*Broyden*)

$f_1(x)=(3-2x_1)x_1-2x_2+1$
$f_n(x)=(3-2x_n)x_n-x_{n-1}+1$
$f_i(x)=(3-2x_i)x_i-x_{i-1}-2x_{i+1}+1,\ i=2, \dots, n-1$
I: $x^0=(-1, \dots, -1)$
II: $10x^0$
III: $100x^0$

Problem 4, $n=4$ (*Powell*)

$f_1(x)=x_1+10x_2$
$f_2(x)=\sqrt{5}(x_3-x_4)$
$f_3(x)=(x_1-2x_3)^2$
$f_4(x)=\sqrt{10}(x_1-x_4)^2$

I: $x^0 = (3, -1, 0, 1)$
II: $10\,x^0$
III: $100\,x^0$

Problem 5, $n=2$ (*Powell*)

$f_1(x) = 10000\,x_1\,x_2 - 1$
$f_2(x) = e^{-x_1} + e^{-x_2} - 1.0001$
I: $x^0 = (0, 1)$
II: $10\,x^0$
III: $100\,x^0$

Problem 6, $n=3$ (*Helical valley*)

$t = \arctan(x_2/x_1)/(2\pi)$
$\theta = t$ if $x_1 \geq 0$ and $\theta = t + 0.5$ if $x_1 < 0$
$f_1(x) = 10(x_3 - 10\,\theta)$
$f_2(x) = 10\,[(x_1^2 + x_2^2)^{1/2} - 1]$
$f_3(x) = x_3$
I: $x^0 = (-1, 0, 0)$
II: $10\,x^0$
III: $100\,x^0$

Problem 7, $n=5$ (*Brown*)

$$f_i(x) = x_i + \sum_{j=1}^{n} x_j - (n+1),\; i = 1, \ldots, n-1$$
$$f_n(x) = \prod_{j=1}^{n} x_j - 1$$
I: $x^0 = (0.5, \ldots, 0.5)$
II: $10\,x^0$
III: $100\,x^0$

Problem 8, $n=3$ (*Brown-Conte*)

$f_1(x) = 3\,x_1 + x_2 + 2\,x_3^2 - 3$
$f_2(x) = -3\,x_1 + 5\,x_2^2 + 2\,x_1\,x_3 - 1$
$f_3(x) = 25\,x_1\,x_2 + 20\,x_3 + 12$
I: $x^0 = (0, 0, 0)$
II: $x^0 = (10, 10, 10)$
III: $x^0 = (100, 100, 100)$

Problem 9, $n=6$ (*Deist-Sefor*)

$$f_i(x) = \sum_{j \neq i} \cot(\beta_i\, x_j)$$

with $\beta=(0.02249, 0.02166, 0.02083, 0.02, 0.01918, 0.01835)$

I: $x^0=(75, \ldots, 75)$

II: $10\,x^0$

III: $100\,x^0$

Problem 10, $n=3$

$f_1(x)=x_1^2+x_2^3+x_3^4-3$

$f_2(x)=\sin(\pi x_1/2)+\cos(\pi x_2/2)+\log(x_3)-1$

$f_3(x)=(1/x_1)+(2/x_2)-(1/x_3)-2$

I: $x^0=(1, 1, 1.5)$

II: $10\,x^0$

III: $100\,x^0$

We tested the new algorithm against the following methods:

(N) Discretized Newton Method

(MN) Modified Discretized Newton Method

(DM) Dennis-Marwil ([5])

(M) Martínez ([15]).

The discretized Newton Method was included only as a point of reference. In fact, the other algorithms were introduced for cases where the application of Newton's method is very expensive which may be not the case of classical test problems. For assuring a fair comparison, the same routines for orthogonal factorizations used in N, MN and the new method, and the same procedures for discretization, stopping criteria and step control were used in all the methods.

The positive integer m has the same meaning for MN, DM, M, and the new method: the method is restarted with a discretized Newton step each time $k\equiv 0 \pmod m$. We used 2 values for m: the optimum value in the sense of Ostrowski ([2, 13, 20]), and $m=2n+1$.

The results are presented in Table 1. The triplet "C, k_1, k_2" means that the method converged in k_1 iterations including k_2 Newton iterations. The triplet D, k_1, k_2 means that the method diverged ($\|F(x^k)\|_\infty \geq 10^8$) with the same meaning for k_1 and k_2. Finally, "$E, 100, k_2$" means that convergence was not achieved after 100 iterations. The experiments were performed in a microcomputer HP-85A, with a machine precision of 2^{-37}. Convergence was accepted when $\|F(x^k)\|_\infty < 2^{-18.5}$.

4. Conclusions

In this paper we have introduced a new Quasi-Newton method for solving nonlinear simultaneous equations. The main feature of this method is that one column of the matrix B_k is changed from one iteration to the next. In the numerical implementation, this column corresponds to the coordinate where the maximum modulus of the step s^k occurs. We obtained a local convergence theorem of the same type as Dennis-Marwil's theorem ([5]).

Table 1

Problem	Case	Newton	Modified Newton $m=\text{opt}$	Modified Newton $m=2n+1$	Dennis-Marwil $m=\text{opt}$	Dennis-Marwil $m=2n+1$	Martínez $m=\text{opt}$	Martínez $m=2n+1$	New $m=\text{opt}$	New $m=2n+1$
1 $n=2$	I	$C,2$	$C,3,1$	$C,3,1$	$C,6,2$	$C,6,2$	$C,9,3$	$C,15,3$	$C,3,1$	$C,3,1$
	II	$C,4$	$C,4,1$	$C,4,1$	$C,7,2$	$C,7,2$	$D,15,4$	$C,11,3$	$C,5,2$	$C,5,1$
	III	$C,4$	$C,5,2$	$C,5,1$	$C,6,2$	$C,6,2$	$D,3,1$	$D,3,1$	$C,6,2$	$C,6,2$
2 $n=2$	I	$C,27$	$D,44,11$	$D,39,8$	$C,78,20$	$C,26,6$	$C,29,8$	$C,26,6$	$C,20,5$	$D,10,2$
	II	$C,23$	$D,36,9$	$D,35,7$	$C,56,14$	$C,41,9$	$C,70,18$	$C,21,5$	$C,73,19$	$C,36,8$
	III	$C,19$	$E,100,25$	$C,73,15$	$C,46,12$	$C,52,11$	$C,69,18$	$C,26,6$	$C,53,14$	$C,69,14$
3 $n=5$	I	$C,4$	$C,7,2$	$C,12,2$	$C,7,2$	$C,9,1$	$C,6,1$	$C,6,1$	$C,7,2$	$C,8,1$
	II	$C,7$	$C,17,3$	$C,23,3$	$C,11,2$	$C,14,2$	$C,9,2$	$C,12,2$	$C,11,2$	$C,14,2$
	III	$C,11$	$C,25,5$	$C,35,4$	$C,15,3$	$C,16,2$	$C,14,3$	$C,14,2$	$C,15,3$	$C,17,2$
4 $n=4$	I	$C,12$	$C,26,6$	$C,37,5$	$C,19,4$	$C,28,4$	$C,16,4$	$C,16,2$	$C,16,4$	$C,16,2$
	II	$C,15$	$C,33,7$	$C,46,6$	$C,24,5$	$C,28,4$	$C,20,4$	$C,21,3$	$C,20,4$	$C,20,3$
	III	$C,18$	$C,41,9$	$C,56,7$	$C,29,6$	$C,56,7$	$C,25,5$	$C,25,3$	$C,25,5$	$C,25,3$
5 $n=2$	I	$C,11$	$C,23,6$	$C,26,6$	$C,15,4$	$C,18,4$	$C,16,4$	$C,16,4$	$C,15,4$	$C,18,4$
	II	$C,4$	$C,13,4$	$C,11,3$	$C,5,2$	$C,6,2$	$C,6,2$	$C,6,2$	$C,5,2$	$C,6,2$
	III	$D,3$	$D,9,3$	$C,46,10$	$D,6,2$	$D,7,2$	$D,2,1$	$D,2,1$	$D,2,1$	$D,2,1$
6 $n=3$	I	$C,9$	$D,30,8$	$D,27,4$	$D,28,7$	$D,40,6$	$D,33,9$	$D,56,8$	$D,34,9$	$D,80,12$
	II	$C,8$	$D,23,6$	$D,22,3$	$D,26,7$	$D,44,7$	$D,87,22$	$D,46,7$	$C,45,12$	$D,27,4$
	III	$C,8$	$D,19,5$	$D,18,3$	$D,22,6$	$D,38,6$	$D,20,5$	$D,20,5$	$D,35,9$	$D,23,4$
7 $n=5$	I	$C,8$	$D,3,1$	$D,10,2$	$C,8,2$	$C,9,1$	$C,10,2$	$C,15,2$	$C,8,2$	$C,9,1$
	II	$C,16$	$D,2,1$	$D,2,1$	$C,21,4$	$C,31,3$	$D,4,1$	$D,4,1$	$E,100,17$	$C,24,3$
	III	$D,1$	$D,1,1$	$D,1,1$	$D,1,1$	$D,1,1$	$D,1,1$	$D,1,1$	$D,1,1$	$D,1,1$
8 $n=3$	I	$C,7$	$D,4,1$	$D,4,1$	$C,10,3$	$C,91,13$	$C,29,8$	$C,15,3$	$C,12,3$	$C,18,3$
	II	$C,8$	$C,18,5$	$C,21,3$	$C,14,4$	$C,14,4$	$C,16,4$	$C,23,4$	$C,17,5$	$C,17,3$
	III	$C,11$	$C,22,6$	$C,36,6$	$C,21,6$	$C,35,5$	$C,20,5$	$C,86,13$	$C,21,6$	$C,19,3$
9 $n=6$	I	$C,6$	$E,100,17$	$E,100,8$	$E,100,17$	$E,100,8$	$C,24,4$	$E,100,8$	$C,9,2$	$C,13,1$
	II	$E,100$	$E,100,17$	$E,100,8$	$E,100,17$	$E,100,8$	$E,100,17$	$E,100,8$	$E,100,17$	$E,100,8$
	III	$E,100$	$E,100,17$	$E,100,8$	$E,100,17$	$E,100,8$	$E,100,17$	$E,100,17$	$E,100,17$	$E,100,8$
10 $n=3$	I	$C,5$	$C,9,3$	$C,10,2$	$C,13,4$	$D,23,4$	$C,19,5$	$C,50,8$	$C,6,2$	$C,6,1$
	II	$D,3$	$D,4,1$	$D,4,1$	$D,11,3$	$D,16,3$	$D,5,2$	$D,15,3$	$D,3,1$	$D,3,1$
	III	$D,3$	$D,4,1$	$D,4,1$	$D,11,3$	$D,16,3$	$D,5,2$	$D,15,3$	$D,3,1$	$D,3,1$

The numerical experiences show that the new method is clearly more efficient than Newton's Modified method. In general, it appears to be also more reliable than Martínez method ([15]) and it seems to be as efficient as Dennis-Marwil's method.

In some cases, we expect that the performance of the new algorithm should be better than the performance of Dennis-Marwil's method. In fact, when the variables of $F(x)$ separate, that is,

$$F(x)=F_1(x_1)+\ldots+F_n(x_n)$$

and we have good initial estimates for some of them, it seems to be useful not to change the columns of $F(x^0)$ corresponding to these variables, as the new method naturally does. This seems to be the reason why the new method outperforms Dennis-Marwil's in the case I of the problems 9 and 10.

As happens to be with Dennis-Marwil's method, the convergence theorem is obtained only for a restarted version of the method. We don't know if a general convergence result for a nonrestarted implementation exists, as in Martínez ([15]) method. However, the numerical experiences show that the convergence properties of these methods are less understood; and we think that much research should be expected in the following years along the lines of [8]. A curious fact is that we have performed a number of numerical experiences with a nonrestarted version of the new method against Broyden's first method (whose convergence is known to be superlinear) and no meaningful differences of efficiency where detected.

As in the case of Dennis-Marwil's and Martínez ([15]) method an efficient version for sparse large scale problems deserves future implementation.

References

[1] Bartels, R. H., Golub, G. H.: The simplex method of linear programming using LU decomposition. Comm. A.C.M. *12*, 266–268 (1969).

[2] Brent, R. P.: Some efficient algorithms for solving systems of nonlinear equations. SIAM J. Numer. Anal. *10*, 327–344 (1973).

[3] Broyden, C. G.: A class of methods for solving nonlinear simultaneous equations. Math Comput. *19*, 577–593 (1965).

[4] Broyden, C. G., Dennis, J. E., Moré, J. J.: On the local and superlinear convergence of Quasi-Newton methods. J. Inst. Math. Appl. *12*, 223–245 (1973).

[5] Dennis, J. E., Marwil, E. S.: Direct secant updates of matrix factorizations. Math. Comput. *38*, 459–476 (1982).

[6] Dennis, J. E., Moré, J. J.: A characterization of superlinear convergence and its application to $Q-N$ methods. Math. Comp. *28*, 549–560 (1974).

[7] Dennis, J. E., Moré, J. J.: Quasi-Newton methods, motivation and theory. SIAM Rev. *19*, 46–89 (1977).

[8] Gay, D. M.: Some convergence properties of Broyden's method. SIAM J. Numer. Anal. *16*, 623–630 (1979).

[9] Gill, P. E., Golub, G., Murray, W., Saunders, M. A.: Methods for modifying matrix factorizations. Math. Comput. *28*, 505–535 (1974).

[10] Goldfarb, D.: On the Bartels-Golub decomposition for linear programming bases. AERE Report C.S.S. 18, Harwell (1975).

[11] Gragg, W. B., Stewart, G. H.: A stable variant of the secant method for solving nonlinear equations. SIAM J. Numer. Anal. *13*, 889–903 (1976).

[12] Martínez, J. M.: Generalization of the methods of Brent and Brown for solving nonlinear simultanous equations. SIAM J. Numer. Anal. *16*, 434–448 (1979).

[13] Martínez, J. M.: Three new algorithms based on the sequential secant method. BIT *19*, 236–243 (1979).

[14] Martínez, J. M.: Solving nonlinear simultaneous equations with a generalization of Brent's method. BIT *20*, 501–510 (1980).

[15] Martínez, J. M.: A Quasi-Newton method with a new updating for the LDU factorization of the approximate Jacobian. Mat. Aplic. e Comput. *2*, 131–142 (1983).
[16] Martínez, J. M., Lopes, T. L.: Combination of the sequential secant method and Broyden's method with projected updates. Computing *25*, 379–386 (1980).
[17] Moré, J. J., Garbow, B. S., Hillstrom, K. E.: Testing unconstrained minimization software ACM TOMS *7*, 17–41 (1981).
[18] Moré, J. J., Tragenstein, J. A.: On the global convergence of Broyden's method. Math. Comp. *30*, 523–540 (1974).
[19] Ortega, J. M., Rheinboldt, W. C.: Iterative Solution of Nonlinear Equations in Several Variables. New York: Academic Press 1970.
[20] Ostrowski, A. M.: Solution of Equations in Euclidean and Banach Spaces. New York-London: Academic Press 1973.
[21] Schmidt, J. W.: Überlinear konvergente Mehrschrittverfahren von Regula-falsi und Newton-Typ. ZAMM *53*, 103–104 (1973).
[22] Schubert, L. K.: Modification of a Quasi-Newton method for nonlinear equations with a sparse Jacobian. Math. Comput. *24*, 27–30 (1970).
[23] Schwetlick, H.: Über die Realisierung und Konvergenz von Mehrschrittverfahren zur iterativen Lösung nichtlinearer Gleichungen. ZAMM *54*, 479–493 (1974).
[24] Schwetlick, H.: Numerische Lösung nichtlinearer Gleichungen. Berlin: Deutscher Verlag der Wissenschaften 1978.

J. M. Martínez
Department of Applied Mathematics
IME CC – UNICAMP
CP 1170
13100 Campinas – SP.
Brazil