ON THE CONVERGENCE OF THE COLUMN-UPDATING METHOD

José Mario Martínez

Setembro RP 34/92

Relatório de Pesquisa

Instituto de Matemática Estatística e Ciência da Computação

UNIVERSIDADE ESTADUAL DE CAMPINAS
Campinas - São Paulo - Brasil

ABSTRACT – The Column-Updating method has proved to be a very efficient technique for solving nonlinear systems of equations. In this paper, we prove new convergence results that tend to explain the numerical behavior of the algorithm.

IMECC – UNICAMP
Universidade Estadual de Campinas
CP 6065
13081 Campinas SP
Brasil

Setembro – 1992

ON THE CONVERGENCE OF THE COLUMN-UPDATING METHOD(*)

José Mario Martínez
Depto. Matemática Aplicada
IMECC - UNICAMP
CP 6065
13081 Campinas SP
Brazil
EMAIL: MARTÍNEZ@CCVAX.UNICAMP.ANSP.BR

Abstract. The Column-Updating method has proved to be a very efficient technique for solving nonlinear systems of equations. In this paper, we prove new convergence results that tend to explain the numerical behavior of the algorithm.
Key words: Nonlinear systems, Quasi-Newton methods, Column-Updating method.

Resumo: O método de Atualização de Coluna tem-se mostrado uma técnica muito eficiente para resolver sistemas não lineares. Neste artigo, provamos novos resultados de convergência que visam explicar o comportamento numérico do algoritmo.
Palavras chave: Sistemas não lineares, métodos quase-Newton, método de Atualização de Coluna.

1.- Introduction

We consider the problem of solving

$$F(x) = 0 \tag{1.1}$$

where $F : \mathbb{R}^n \to \mathbb{R}^n$ is differentiable. The Newton method, given by the iteration

$$x_{k+1} = x_k - [F'(x_k)]^{-1}F(x_k), \tag{1.2}$$

(*) This work was supported by FAPESP (Projeto Temático 90-3724-6).

is the most popular technique for solving (1.1). See Ortega and Rheinboldt [1970], Dennis and Schnabel [1983], Schwetlick [1978], Ostrowski [1973], etc.. When the derivatives of F are not available, or when the cost of solving the linear system associated with (1.2) is prohibitive, the more general iteration

$$x_{k+1} = x_k - B_k^{-1} F(x_k) \tag{1.3}$$

is used, where B_k is an approximation of the Jacobian which only uses available information. In general, the resolution of the system associated with (1.3) is less expensive than (1.2) due to suitable updating procedures. The methods based on (1.3) are called quasi-Newton methods. See Dennis and Moré [1977], Dennis and Schnabel [1983].

The best known class of quasi-Newton methods is the class of Least Change Secant Update (LCSU) algorithms (Dennis and Schnabel [1979, 1983]). For this class of methods, under suitable assumptions, Q-superlinear convergence can be proved, provided that x_0 and B_0 are good approximations of the solution x_* and the Jacobian $F'(x_*)$ respectively. See Dennis and Walker [1981], Martínez [1990, 1992a].

However, some authors have introduced quasi-Newton methods that don't belong to the LCSU family, but seem to be useful in practice. See Gomes-Ruggiero, Martínez and Moretti [1992]. This is the case of the Column-Updating Method, studied in this paper. See Martínez [1984a], Gomes-Ruggiero and Martínez [1992]. In the Column-Updating Method B_{k+1} is obtained from B_k by changing only one suitable column so that a "secant" equation is satisfied. The numerical results obtained with this method both in small and large systems of equations are very good, but the theoretical convergence results are weak.

In this paper, we try to fill the gap between theory and practice, in relation to the Column-Updating Method. Briefly speaking, we prove that, under classical assumptions, linear convergence implies R-superlinear convergence (Section 2) and that for a restarted version of the algorithm (not necessarily Newton restarts) local, linear and, thus, R-superlinear convergence takes place (Section 3). Moreover, we prove in Section 4 that stronger results are true in the two-dimensional case.

In the proofs of Section 2 we made a strong utilization of the results of Gay [1979] concerning finite convergence of rank-one secant methods for linear systems. In the proof of Theorem 2.1 we adopted the O(.) notation in order to simplify the exposition. A careful reader can easily verify that the constants involved in the bounds are independent of k. Our experience in reading and writing convergence proofs taught us that proofs where all the constants are exhaustively defined are much more difficult to read and, many times, less convincing, than those where the O(.) notation is used.

2.- The general algorithm without restarts

We will consider the problem of solving (1.1), where $F : \Omega \subset I\!R^n \to I\!R^n$, Ω is an open and convex set and $F \in C^1(\Omega)$. We denote $J(x)$ the Jacobian matrix of $F(x)$. Let $||\cdot||$ be an arbitrary norm on $I\!R^n$. We assume that there exists $L > 0$, $x_* \in \Omega$ such that

$$||J(x) - J(x_*)|| \leq L||x - x_*|| \tag{2.1}$$

for all $x \in \Omega$. This implies (see Broyden, Dennis and Moré [1973]) that

$$||F(z) - F(x) - J(x_*)(z - x)|| \leq L||z - x||\sigma(x, z) \tag{2.2}$$

for all $x, z \in \Omega$, where $\sigma(x, z) = \max\{||x - x_*||,\ ||z - x_*||\}$. Moreover, we assume that $F(x_*) = 0$ and $J(x_*)$ is nonsingular.

Let us define the main algorithm considered in this section. This algorithm is a generalization of the Column-Updating method and, in fact, it also generalizes the classical Broyden's [1965] method.

Algorithm 2.1

Let $x_0 \in \Omega$, B_0 an arbitrary nonsingular initial matrix, $\alpha \in (0, 1]$. For all $k = 0, 1, 2, \ldots$ such that $F(x_k) \neq 0$, perform the following steps.

Step 1. Compute

$$s_k = -B_k^{-1}F(x_k), \tag{2.3}$$

$$x_{k+1} = x_k + s_k,$$

$$y_k = F(x_{k+1}) - F(x_k). \tag{2.4}$$

Step 2. Choose $u_k \in I\!R^n$ such that $||u_k|| = 1$,

$$|u_k^T s_k| \geq \alpha||s_k|| \tag{2.5}$$

and

$$u_k^T B_k^{-1} y_k \neq 0 \tag{2.6}$$

Step 3. Define

$$B_{k+1} = B_k + \frac{F(x_{k+1})u_k^T}{u_k^T s_k}. \tag{2.7}$$

Remarks

In the Column-Updating method we choose u_k as a canonical vector that satisfies (2.5).

Applying the Sherman-Morrison-Woodbury formula (see Golub and Van Loan [1989, page 51]) we verify that B_{k+1} is nonsingular if (2.6) holds. In this section we will assume that a choice of u_k satisfying (2.5) and (2.6) is always possible.

Assumption A1

Let us assume that an infinite sequence generated by Algorithm 2.1 is well defined and that the following statements are true:

a) $\lim_{k\to\infty} x_k = x_*$. (2.8)

b) There exists $r \in (0,1)$ such that

$$||x_{k+1} - x_*|| \leq r||x_k - x_*|| \tag{2.9}$$

for all $k = 0, 1, 2, \ldots$.

c) There exists $M > 0$ such that

$$||B_k|| \leq M, \ \ ||B_k^{-1}|| \leq M \tag{2.10}$$

for all $k = 0, 1, 2, \ldots$.

Under Assumption A1 we will prove now that the convergence is R-superlinear.

Theorem 2.1

Consider Algorithm 2.1 and suppose that Assumption A1 is satisfied. Then

$$\lim_{k\to 0} \frac{||x_{k+2n} - x_*||}{||x_k - x_*||} = 0. \tag{2.11}$$

Proof. Suppose that (2.11) is not true. Then there exists K_1, an infinite subset of $\mathbb{N}$, $c > 0$ such that

$$||x_{k+2n} - x_*|| \geq c||x_k - x_*|| \tag{2.12}$$

for all $k \in K_1$.

Let $K_1 = \{k_1, k_2, \ldots\}, k_1 < k_2 < \ldots$. We assume, without loss of generality, that $k_{i+1} \geq k_i + 2n$ for all $i = 1, 2, 3, \ldots$. Define

$$\varepsilon_k = ||x_k - x_*||, \; k = 0, 1, 2, \ldots. \tag{2.13}$$

By (2.5) and (2.9), we have that

$$|u_k^T s_k| \geq c_1 \varepsilon_k \tag{2.14}$$

for all $k = 0, 1, 2, \ldots$, where $c_1 = \alpha(1 - r)$. Moreover, by (2.9), (2.12) and (2.13),

$$\varepsilon_{k+\ell} \geq c\varepsilon_k \tag{2.15}$$

for all $k \in K_1$, $\ell \in \{0, 1, \ldots, 2n\}$.

For all $k \in K_1$, we define $x_{k,\ell}, \ell = 0, 1, \ldots, 2n$, by:

$$\text{(a) } x_{k,0} = x_k, \; B_{k,0} = B_k, \tag{2.16}$$

$$\text{(b) } s_{k,\ell} = -B_{k,\ell}^{-1} A(x_{k,\ell} - x_*), \tag{2.17}$$

$$\text{(c) } x_{k,\ell+1} = x_{k,\ell} + s_{k,\ell}, \tag{2.18}$$

$$\text{(d) } B_{k,\ell+1} = B_{k,\ell} + \frac{A(x_{k,\ell+1} - x_*)u_{k+\ell}^T}{u_{k+\ell}^T s_{k,\ell}}, \tag{2.19}$$

where $A = J(x_*)$. Clearly, $x_{k,\ell+1}$ will be well defined for all $\ell = 0, 1, \ldots, 2n - 1$ if $u_{k+\ell}^T s_{k,\ell} \neq 0$ and $B_{k,\ell}$ is nonsingular. We will see that this is the case if k is large enough. More precisely, we will prove that, for large enough $k \in K_1$, $x_{k,1}, x_{k,2}, \ldots, x_{k,2n}$ are well defined, $B_{k,1}, B_{k,2}, \ldots, B_{k,2n-1}$ are well defined and nonsingular, and

$$\text{(a) } ||x_{k+\ell} - x_{k,\ell}|| = O(\varepsilon_k^2), \tag{2.20}$$

$$\text{(b) } ||B_{k,\ell} - B_{k+\ell}|| = O(\varepsilon_k), \tag{2.21}$$

$$\text{(c) } ||B_{k,\ell}^{-1} - B_{k+\ell}^{-1}|| = O(\varepsilon_k). \tag{2.22}$$

Let us proceed by induction on ℓ. For $\ell = 0$, (2.20)-(2.22) are trivial. Assume that (2.20)-(2.22) hold for some fixed ℓ. Then,

$$||x_{k+\ell+1} - x_{k,\ell+1}|| = ||x_{k+\ell} - B_{k+\ell}^{-1} F(x_{k+\ell}) - x_{k,\ell} - B_{k,\ell}^{-1} A(x_{k,\ell} - x_*)||$$

$$\leq ||x_{k+\ell} - x_{k,\ell}|| + ||B_{k+\ell}^{-1}|| \, ||F(x_{k+\ell}) - A(x_{k,\ell} - x_*)||$$

$$+ ||B_{k+\ell}^{-1} - B_{k,\ell}^{-1}|| \, ||A(x_{k,\ell} - x_*)||. \tag{2.23}$$

But, by the inductive hypothesis and (2.2),

$$||F(x_{k+\ell}) - A(x_{k,\ell} - x_*)|| \leq ||F(x_{k+\ell}) - A(x_{k+\ell} - x_*)|| + ||A|| \, ||x_{k+\ell} - x_{k,\ell}|| = O(\varepsilon_k^2). \tag{2.24}$$

So, by (2.23), (2.24) and the inductive hypothesis,

$$||x_{k+\ell+1} - x_{k,\ell+1}|| = O(\varepsilon_k^2). \tag{2.25}$$

By (2.25) and the inductive hypothesis we deduce that

$$||s_{k+\ell} - s_{k,\ell}|| = O(\varepsilon_k^2). \tag{2.26}$$

Now, by (2.14) and (2.26),

$$\begin{aligned} |u_{k+\ell}^T \, s_{k,\ell}| &= |u_{k+\ell}^T s_{k+\ell} + u_{k+\ell}^T (s_{k,\ell} - s_{k+\ell})| \\ &\geq c_1 \varepsilon_{k+\ell} - O(\varepsilon_k^2) \\ &\geq c_1 c \varepsilon_k - O(\varepsilon_k^2) \geq c_2 \varepsilon_k > 0 \end{aligned} \tag{2.27}$$

if k is large enough, where $c_2 = c_1 c/2$.

By (2.27), $B_{k,\ell+1}$ is well defined. So, by the inductive hypothesis,

$$||B_{k,\ell+1} - B_{k+\ell+1} \leq O(\varepsilon_k) + \left\| \frac{F(x_{k+\ell+1})}{u_{k+\ell}^T \, s_{k+\ell}} - \frac{A(x_{k,\ell+1} - x_*)}{u_{k+\ell}^T s_{k,\ell}} \right\| \tag{2.28}$$

Now,

$$\begin{aligned} &\left\| \frac{F(x_{k+\ell+1})}{u_{k+\ell}^T \, s_{k+\ell}} - \frac{A(x_{k,\ell+1} - x_*)}{u_{k+\ell}^T s_{k,\ell}} \right\| \\ &\leq \left\| \frac{F(x_{k+\ell+1}) - A(x_{k,\ell+1} - x_*)}{u_{k+\ell}^T s_{k+\ell}} \right\| \\ &+ \left\| \frac{A(x_{k+\ell+1} - x_{k,\ell+1})}{u_{k+\ell}^T s_{k+\ell}} \right\| \\ &+ \left\| A(x_{k,\ell+1} - x_*) \left(\frac{1}{u_{k+\ell}^T s_{k+\ell}} - \frac{1}{u_{k+\ell}^T s_{k,\ell}} \right) \right\|. \end{aligned} \tag{2.29}$$

By (2.2), (2.27) and (2.9), we have that

$$\left\|\frac{F(x_{k+\ell+1}) - A(x_{k,\ell+1} - x_*)}{u_{k+\ell}^T s_{k+\ell}}\right\| \leq \frac{M\|x_{k+\ell+1} - x_*\|^2}{c_2 \varepsilon_k} \leq \frac{M r^{\ell+1}}{c_2}\varepsilon_k = O(\varepsilon_k). \tag{2.30}$$

By (2.25) and (2.27),

$$\frac{\|A(x_{k+\ell+1} - x_{k,\ell+1})\|}{|u_{k+\ell}^T s_{k+\ell}|} \leq \frac{\|A\| O(\varepsilon_k^2)}{c_2 \varepsilon_k} = O(\varepsilon_k). \tag{2.31}$$

By (2.26), (2.14) and (2.27),

$$\left|\frac{1}{u_{k+\ell}^T s_{k+\ell}} - \frac{1}{u_{k+\ell}^T s_{k,\ell}}\right| = \left|\frac{u_{k+\ell}^T (s_{k,\ell} - s_{k+\ell})}{u_{k+\ell}^T s_{k+\ell} u_{k+\ell}^T s_{k,\ell}}\right| \leq \frac{O(\|s_{k,\ell} - s_{k+\ell}\|)}{c_1 c_2 \varepsilon_k^2} = O(1). \tag{2.32}$$

Therefore, by (2.9) and (2.25),

$$\left\|A(x_{k,\ell+1} - x_*)\left(\frac{1}{u_{k+\ell}^T s_{k+\ell}} - \frac{1}{u_{k+\ell}^T s_{k,\ell}}\right)\right\| = O(\varepsilon_k). \tag{2.33}$$

By (2.28), (2.30), (2.31) and (2.33) we have that

$$\|B_{k,\ell+1} - B_{k+\ell+1}\| = O(\varepsilon_k). \tag{2.34}$$

Since $\|B_k^{-1}\|$ is bounded, we obtain, using Banach's perturbation lemma (See Golub and Van Loan [1989, pp. 59-60]) that $B_{k,\ell+1}^{-1}$ exists and

$$\|B_{k,\ell+1}^{-1} - B_{k+\ell+1}^{-1}\| = O(\varepsilon_k). \tag{2.35}$$

if k is large enough.

Now, by Theorem 2.2 of Gay [1979], we have that $x_{k,2n} = x_*$ for all $k \in K_1$ such that $x_{k,2n}$ is well defined. So, by (2.20) with $\ell = 2n$, we have that $\varepsilon_{k+2n} = O(\varepsilon_k^2)$ if $k \in K_1$. This contradicts (2.15).

Therefore, the set K_1 cannot exist and, thus, (2.11) is proved. □

Corollary 2.1

If the hypotheses of Theorem 2.1 are satisfied then the convergence of x_k to x_* is R-superlinear.

Proof. We need to prove that

$$\lim_{k\to\infty} \varepsilon_k^{1/k} = 0. \tag{2.36}$$

See Ortega and Rheinboldt [1970, Section 9.2].

By Theorem 2.1 we have that

$$\lim_{p\to\infty} \frac{\varepsilon_{2n(p+1)}}{\varepsilon_{2np}} = 0. \tag{2.37}$$

So, the sequence $\gamma_p \equiv \varepsilon_{2np}$ converges Q-superlinearly to 0. Therefore, γ_p converges R-superlinearly to 0. So,

$$\lim_{p\to\infty} \varepsilon_{2np}^{1/p} = 0. \tag{2.38}$$

Clearly, by (2.38),

$$\lim_{p\to\infty} \varepsilon_{2np}^{1/2np} = 0. \tag{2.39}$$

So, if $1 \leq \ell \leq 2n$, we have, by (2.9) and (2.39), that

$$\lim_{p\to\infty} \varepsilon_{2np+\ell}^{1/(2np+\ell)} \leq \lim_{p\to\infty} \left(\varepsilon_{2np}^{1/2np}\right)^{\frac{2np}{2np+\ell}} = 0,$$

since $\dfrac{2np}{2np+\ell} \to 1$. This completes the proof.□

3 - The restarted algorithm

In this section we analyze a variation of Algorithm 2.1 that consists of computing x_{k+1} and B_{k+1} as in Algorithm 2.1, except when $k+1$ is a multiple of a fixed integer q. In this case, we set $B_{k+1} = C(x_{k+1})$, where $C(x)$ is some approximation of $J(x)$ for all $x \in \Omega$. Usually $C(x)$ is the projection of $J(x)$ on a suitable subspace, or it represents an incomplete factorization of $J(x)$.

Algorithm 3.1

Let $x_o \in \Omega, B_o = C(x_o), \alpha \in (0,1], q$ a fixed integer. If $F(x_k) \neq 0$, compute s_k and x_{k+1} as in Step 1 of Algorithm 2.1. If $k+1$ is not a multiple of q, compute u_k and B_{k+1} performing steps 2 and 3 of Algorithm 2.1. If $k+1$ is a multiple of q, set

$$B_{k+1} = C(x_{k+1}). \tag{3.1}$$

As in the case of Algorithm 2.1, the Column-Updating method corresponds to the following choice of u_k:

$$u_k = e_{j_k}$$

where e_{j_k} is a canonical vector such that (2.5) holds. In general, we choose

$$j_k = \text{Argmax } \{|e_j^T s_k|, j = 1, \ldots, n\}. \tag{3.2}$$

Theorem 3.1

Let $r \in (0,1)$. There exist $\varepsilon, \delta > 0$ such that, if $\|x_o - x_*\| \leq \varepsilon$ and $\|C(x_k) - J(x_*)\| \leq \delta$ for all $k \equiv 0 (\text{mod } q)$ then the sequence generated by Algorithm 3.1 is well defined, converges to x_* and satisfies

$$\|x_{k+1} - x_*\| \leq r\|x_k - x_*\| \tag{3.3}$$

for all $k = 0, 1, 2, \ldots$. Moreover, $\|B_k\|$ and $\|B_k^{-1}\|$ are bounded.

Proof. See the proof of Theorem 3.1 of Gomes-Ruggiero and Martínez [1992]. The adaptation to our case involves two differences: first, here we are dealing with the general form of u_k while in that paper $u_k \equiv e_{j_k}$. However, this difference does not influentiate the proof. Second, we are not assuming here that $B_k = J(x_k)$ when k is a multiple of q. For this reason we only obtain (3.3) instead of Q-superlinear convergence.□

By Theorem 3.1, we know that, under reasonable conditions, we can obtain linear convergence and boundedness of $\|B_k\|$and $\|B_k^{-1}\|$ for Algorithm 3.1. Remember that this is not the case of the non-restarted Algorithm 2.1, for which we don't have reasonable sufficient conditions for linear convergence.

Theorem 3.2

Consider Algorithm 3.1, with $q \geq 2n$. Assume that a well defined infinite sequence is generated, $\|B_k\|$ and $\|B_k^{-1}\|$ are bounded and (3.3) holds. Then

$$\lim_{j \to \infty} \frac{\|x_{jq+2n} - x_*\|}{\|x_{jq} - x_*\|} = 0. \tag{3.4}$$

Moreover, the convergence of (x_k) to x_* is $R-$superlinear.

Proof. For proving (3.4) proceed by contradiction in the same way as we did in Theorem 2.1. The $R-$superlinear convergence follows as in Corollary 2.1.□

4 - The Column - Updating Method without restarts when $n = 2$

In this section we consider the particular implementation of Algorithm 2.1 that corresponds to the Column-Updating Method as given in Martínez [1984a] and Gomes-Ruggiero and Martínez [1992]. So, we choose

$$u_k = e_{j_k} \tag{4.1}$$

where

$$j_k = \underset{j}{\text{Argmax}} \ \{|e_j^T s_k|\} \tag{4.2}$$

and $\{e_1, \ldots, e_n\}$ is the canonical basis of $I\!R^n$. We will prove that in the two-dimensional case, with the choices (4.1) and (4.2), the assumption A1 holds if x_o and B_o are close enough to x_* and $J(x_*)$ respectively. This result is based essentially on the following "bounded deterioration" property.

Lemma 4.1

Assume that $n = 2$. Consider Algorithm 2.1 with the choices (4.1) and (4.2). Suppose that $x_o, x_1, \ldots, x_{k+1}$ are well defined . Then

$$\|B_{k+1} - J(x_*)\|_1 \leq \|B_k - J(x_*)\|_1 + \frac{L}{2}\sigma(x_k, x_{k+1}). \tag{4.3}$$

Proof. By (2.2), (2.4) and (2.7), we have that

$$\begin{aligned}\|B_{k+1} - J(x_*)\|_1 &= \|B_k + \frac{(y_k - B_k s_k)e_{j_k}^T}{e_{j_k}^T s_k} - J(x_*)\|_1 \\ &\leq \|B_k + \frac{(J(x_*) - B_k)s_k e_{j_k}^T}{e_{j_k}^T s_k} - J(x_*)\|_1 \\ &+ \|\frac{(y_k - J(x_*)s_k)e_{j_k}^T}{e_{j_k}^T s_k}\|_1.\end{aligned} \tag{4.4}$$

Now, by (2.2) and (4.2),

$$\begin{aligned}\|\frac{(y_k - J(x_*)s_k)e_{j_k}^T}{e_{j_k}^T s_k}\|_1 &\leq \frac{L\|s_k\|_1 \ \sigma(x_k, x_{k+1})}{\|s_k\|_1/2} \\ &\leq 2L \ \sigma(x_k, x_{k+1}).\end{aligned} \tag{4.5}$$

Moreover,

$$\|B_k + \frac{(J(x_*) - B_k)s_k e_{j_k}^T}{e_{j_k}^T s_k} - J(x_*)\|_1$$
$$\leq \|B_k - J(x_*)\|_1 \quad \|I - \frac{s_k e_{j_k}^T}{e_{j_k}^T s_k}\|_1. \tag{4.6}$$

Now, by (4.1) and (4.2), if $j_k = 1$ we have

$$I - \frac{s_k e_{j_k}^T}{e_{j_k}^T s_k} = \begin{pmatrix} 0 & 0 \\ -\frac{e_2^T s_k}{e_1^T s_k} & 1 \end{pmatrix}. \tag{4.7}$$

If $j_k = 2$, we have that

$$I - \frac{s_k e_{j_k}^T}{e_{j_k}^T s_k} = \begin{pmatrix} 1 & \frac{e_1^T s_k}{e_2^T s_k} \\ 0 & 0 \end{pmatrix}. \tag{4.8}$$

So, by (4.2), (4.7) and (4.8),

$$\|I - \frac{s_k e_{j_k}^T}{e_{j_k}^T s_k}\|_1 = 1. \tag{4.9}$$

Thus, (4.3) follows from (4.4), (4.5), (4.6) and (4.9).□

Theorem 4.1

Let $r \in (0,1), n = 2$. There exist $\varepsilon, \delta > 0$ such that, if $\|x_o - x_*\| \leq \varepsilon, \|B_o - J(x_*)\| \leq \delta$, the sequence generated by Algorithm 2.1 with the choices (4.1) - (4.2) is well defined, converges to x_* and satisfies

$$\|x_{k+1} - x_*\|_1 \leq r\|x_k - x_*\|_1 \tag{4.10}$$

for all $k = 0, 1, 2, \ldots$. Moreover, $\|B_k\|$ and $\|B_k^{-1}\|$ are bounded and for any norm $\|\cdot\|$, (2.11) holds and the convergence is $R-$superlinear.

Proof. (4.10) and the boundedness of $\|B_k\|, \|B_k^{-1}\|$ follow from (4.3) and the general assumptions of section 2 using a classical inductive proof. See, for example, the proof of Theorem 3.2 of Broyden, Dennis and Moré [1973]. Therefore, by Theorem 2.1, (2.11) holds for $\|\cdot\|_1$. Hence, (2.11) holds for any norm since superlinear convergence is norm-independent. Finally, the convergence is $R-$superlinear by Corollary 2.1.□

5 - Conclusions

Since its introduction in 1984, the excellent numerical performance of the Column-Updating method (Martínez [1984 a,1984 b], Gomes-Ruggiero and Martínez [1992], Martínez and Zambaldi [1991]) has been intriguing. In fact, the method has shown to be comparable, and many times superior, to Broyden's method (Broyden [1965], Dennis and Schnabel [1983]), not only in terms of global computational effort, but also in terms of robustness and number of iterations. However, it is clear that the general local convergence theories that guarantee local and superlinear convergence of quasi-Newton methods (Dennis and Walker [1981], Martínez [1990,1992 a, 1992 b]) are not applicable to the Column-Updating method. Moreover, the convergence properties that were proved in Martínez [1984a] and Gomes-Ruggiero and Martínez [1992] are the same (in fact, slightly weaker) as those that can be proved for the Modified Newton method, where the Jacobian approximation is not modified at all throughout the process.

In this paper we proved some theorems that tend to explain why the Column-Updating Method is so good. Briefly speaking, we proved that:
(a) If linear convergence is assumed, the convergence is $R-$superlinear.
(b) If the method is restarted periodically (not necessarily with true Jacobians) we obtain local and $R-$superlinear convergence.
(c) If $n = 2$, no restarts are necessary for proving local and $R-$ superlinear convergence.

Of course, these results are still weaker than the properties that hold for Broyden's method and other Least-Change - Secant Update methods but, at least, are much stronger than the convergence properties of the Modified Newton method.

Using the techniques introduced here, there same results can be proved for the Inverse-Column-Updating method introduced by Martínez and Zambaldi [1992].

References

C.G. Broyden [1965] : A class of methods for solving nonlinear simultaneous equations, **Mathematics of Computation 19**, pp. 577-593.

C.G. Broyden, J. E. Dennis and J. J. Moré [1973]: On the local and superlinear convergence of quasi-Newton methods, **Journal of the Institute of Mathematics and Applications 12**, pp. 223-245.

J. E. Dennis and J.J. Moré [1977]: Quasi-Newton methods, motivation and theory,

SIAM Review 19, pp. 46-89.

J.E. Dennis and R.B. Schnabel [1979]: Least - Change Secant Updates for quasi-Newton methods, **SIAM Review 21**, pp. 443-459.

J.E. Dennis and R.B. Schnabel [1983]: **Numerical Methods Unconstrained Optimization and Nonlinear Equations**, Prentice Hall, Englewood Cliffs, NJ.

J.E. Dennis and H.F. Walker [1981]: Convergence theorems for Least-Change Secant Update methods, **SIAM Journal on Numerical Analysis 18**, pp. 949-987.

D.M. Gay [1979]: Some convergence properties of Broyden's methhod, **SIAM Journal on Numerical Analysis 16**, pp. 623-630.

G.H. Golub and Ch. F. Van Loan [1989]: **Matrix Computations**, The Johns Hopkins University Press, Baltimore and London.

M.A. Gomes-Ruggiero and J. M. Martínez [1992]: The Column-Updating method for solving nonlinear equations in Hilbert space, **Mathematical Modelling and Numerical Analysis 26**, pp. 309-330.

M.A. Gomes-Ruggiero, J. M. Martínez and A.C. Moretti [1992]: Comparing algorithms for solving sparse nonlinear systems of equations, to appear in **SIAM Journal on Scientific and Statistical Computing**.

J.M. Martínez [1984 a]: A quasi-Newton method with modification of one column per iteration, **Computing 33**, pp. 353-362.

J.M. Martínez [1984 b]: **Métodos quase-Newton com atualização direta das fatorações**, tese de Livre-Docência em Análise Numérica, UNICAMP.

J.M. Martínez [1990]: Local Convergence Theory of Inexact Newton methods based on structured Least-Change updates, **Mathematics of Computation 55**, pp. 143-168.

J.M.Martínez [1992 a]: On the relation between two local convergence theories of Least-Change Secant Update methods, to appear in **Mathematics of Computation**.

J.M.Martínez [1992 b]: Fixed-Point quasi-Newton methods, to appear in **SIAM Journal on Numerical Analysis.**

J.M.Martínez and M.C. Zambaldi [1992]: An Inverse Column-Updating method for solving large-scale nonlinear systems of equations, to appear in **Optimization Methods and Software.**

J.M.Ortega and W.C. Rheinboldt [1970]: **Iterative Solution of Nonlinear Equations in Several variables**, Academic Press, New York .

A.M.Ostrowski [1973]: **Solution of Equations in Euclidean and Banach Spaces**, Academic Press, New York.

H. Schwetlick [1978]: **Numerische Lösung nichtlinearer Gleichungen**, Deutscher Verlag der Wissenschaften, Berlin.

RELATÓRIOS DE PESQUISA — 1992

01/92 Uniform Approximation the: Non-locally Convex Case — *João B. Prolla.*

02/92 Compactificação de $L^{\tau}_{\omega\omega}(Q)$ com τ Finito — *A. M. Sette* and *J. C. Cifuentes.*

03/92 Um Modelo para Aquisição da Especificação — *Cecilia Inés Sosa Arias* and *Ariadne Carvalho.*

04/92 Convergence Estimates for the Wavelet Galerkin Method — *Sônia M. Gomes* and *Elsa Cortina.*

05/92 Optimal Chemotherapy: A Case Study with Drug Resistance, Saturation Effect and Toxicity — *M. I. S. Costa, J. L. Boldrini* and *R. C. Bassanezi.*

06/92 On the Paper "Cauchy Completeness of Elementary Logic" of D. Mundici and A. M. Sette — *J. C. Cifuentes.*

07/92 What is the EM Algorithm for Maximum Likelihood Estimation in PET and How to Accelerate it — *Alvaro R. De Pierro.*

08/92 Bifurcation from infinity and multiple solutions for an elliptic system — *Raffaele Chiappinelli* and *Djairo G. de Figueiredo.*

09/92 Approximation Processes for Vector-Valued Continuous Functions — *João B. Prolla.*

10/92 Aplicação do Método de Fraissé à Compactificação de Lógicas com Quantificadores Co-filtro — *A. M. Sette* and *J. C. Cifuentes.*

11/92 Absolutely Summing Holomorphic Mappings — *Mário C. Matos.*

12/92 The Feynman-Dyson Proof of Maxwell Equations and Magnetic Monopoles — *Adolfo A. Jr.* and *Waldyr A. R. Jr.*

13/92 A Generalized Dirac's Quantization Condition for Phenomenological Non-abelian Magnetic Monopoles — *Adolfo M. Jr.* and *Waldyr A. R. Jr.*

14/92 Multiplicity Results for the 1-Dimensional Generalized p-Laplacian — *Pedro Ubilla.*

15/92 Nowhere Vanishing Torsion Closed Curves Always Hide Twice — *Sueli R. Costa.* and *Maria Del Carmen R. Fuster.*

16/92 Uniform Approximation of Continuous Convex–Cone–Valued Functions — *João B. Prolla.*

17/92 Monotonically Dominated Operators on Convex Cones — *A. O. Chiacchio, J. B. Prolla, M. L. B. Queiroz* and *M. S. M. Roversi.*

18/92 Testing the Concept of a Photon as an Extended Object in a Variation of Franson's Experiment — *V. Buonomano, A. J. R. Madureira* and *L. C. B. Ryff.*

19/92 A New Trust Region Algorithm for Boun Constrained Minimization — *Ana Friedlander, José Mario Martínez* and *Sandra A. Santos.*

20/92 **A Priori Estimates for Positive Solutions of Semilinear Elliptic Systems Via Hardy-Sobolev Inequalities** — *Ph. Clément, D.G. de Figueiredo* and *E. Mitidieri.*

21/92 **On the Resolution of Linearly Constrained Convex Minimization Problems** — *Ana Friedlander, José Mario Martínez* and *Sandra A. Santos.*

22/92 **Convergência no Espectro Real de um Anel** — *J. C. Cifuentes.*

23/92 **Parallel Methods for Large Least Norm Problems** — *Alvaro R. De Pierro.* and *Jose M. Lopes.*

24/92 **A Generalization of Dirac Non–Linear Electrodynamics, and Spinning Charged Particles** — *Waldyr A. Rodrigues Jr., Jayme Vaz Jr.* and *Erasmo Recami.*

25/92 **Bifurcation of Certain Singularities of Sliding Vector Fields** — *Marco Antonio Teixeira.*

26/92 **Density of Infimum-Stable Convex Cone** — *João B. Prolla.*

27/92 **An Extended Decomposition Through Formalization in Product Spaces** — *Alvaro R. De Pierro.*

28/92 **Accelerating Iterative Algorithms for Symmetric Linear Complementarity Problems** — *Alvaro R. De Pierro* and *José Marcos Lópes.*

29/92 **Free Maxwell Equations, Dirac Equation, and Non-Dispersive de Broglie Wave–packets** — *Waldyr A. Rodrigues Jr., Jayme Vaz Jr.* and *Erasmo Recami.*

30/92 **New Results on the Equivalence of Regularization and Truncated Iteration for Ill-posed Problems** — *Reginaldo J. Santos.*

31/92 **Topological Defects: A Distribution Theory Approach** — *Patricio S. Letelier* and *Anzhong Wang.*

32/92 **Some Applications of the Generalized Khintchine's Inequality** — *Mário C. Matos.*

33/92 **A Technique for Finding The Center of A Polytope** — *Antonio Carlos Moretti.*