Journal of Computational Mathematics, Vol.11, No.3, 1993, 236–249.

ON THE CONVERGENCE OF THE FACTORIZATION UPDATE ALGORITHM*

Bai Zhong-zhi   Wang De-ren
(*Shanghai University of Science and Technology, Shanghai, China*)

Abstract

In this paper, we make a Kantorovich-type analysis for the sparse Johnson and Austria's algorithm given in [2], which is called factorization update algorithm. When the mapping is linear, it is shown that a modification of that algorithm leads to global and Q-superlinear convergence. Finally, we point out the modification is also of local and Q-superlinear convergence for nonlinear systems of equations and give its corresponding Kantorovich-type convergence result.

§1. Introduction

For a large sparse nonlinear system of equations

$$F(x)=0, \qquad F: D\subset R^n \longrightarrow R^n, \tag{1.1}$$

Johnson and Austria gave a kind of direct secant updates employing matrix factorizations to get its solution and proved its local and Q-superlinear convergence property(see [1]). By changing the updates in matrix forms, the authors[3] set up its Kantorovich-type analysis. Since Johnson and Austria's algorithm does not maintain the sparse structures of the triangular factors, the authors proposed its modified version, obtained a factorization update algorithm which has the sparse transitivity property of triangular factors H and U, and proved that this algorithm is of local and Q-superlinear convergence (see [2]).

In this paper, in order to complete the convergence theory of the factorization update algorithm, we make a Kantorovich-type analysis paralleling the one given in [3]. When $F(x)$ is linear, the Kantorovich-type convergence theorem naturally leads to global and Q-superlinear convergence for the modified factorization update algorithm. Finally, we show that the modified algorithm is convergent locally and Q-superlinearly for solving the large sparse nonlinear system of equations (1.1).

In the remainder of this section, we restate the factorization update algorithm for reference.

First, we introduce the following notations given in [2]:

* Received October 8, 1991.

Set $P_n \subset R^n \times R^n, P_n = \{(i,j) | i \neq j, 1 \leq i,j \leq n\}$; for certain $P \subset P_n$, subspace $\mathcal{L} \subset R^n \times R^n, \mathcal{L} = \{L \in L(R^n) | L$ is a unit lower triangular matrix, $L_{ij} = 0$, for $(i,j) \in P$ and $i > j, 1 \leq i,j \leq n\}$; subspace $\mathcal{U} \subset R^n \times R^n$, $\mathcal{U} = \{U \in L(R^n) | U$ is an upper triangular matrix, $U_{ij} = 0$ for $(i,j) \in P$ and $i \leq j, 1 \leq i,j \leq n\}$; matrices $\Lambda_1 = (\lambda_{ij}^{(1)}), \Lambda_2 = (\lambda_{ij}^{(2)}) \in L(R^n)$,

$$\lambda_{ij}^{(1)} = \begin{cases} 1, & \text{if } U_{ij} \neq 0, \\ 0, & \text{otherwise}, \end{cases} \qquad \lambda_{ij}^{(2)} = \begin{cases} 1, & \text{if } H_{ij} \neq 0, \text{ and } i > j, \\ 0, & \text{otherwise} \end{cases}$$

and n-dimensional vectors

$$s(i)^T = (s(i)_1, s(i)_2, \cdots, s(i)_n),$$
$$y(i)^T = (y(i)_1, y(i)_2, \cdots, y(i)_n),$$

with

$$s(i)_j = \begin{cases} s_j, & \text{if } \lambda_{ij}^{(1)} = 1, \\ 0, & \text{otherwise}, \end{cases} \qquad y(i)_j = \begin{cases} y_j, & \text{if } \lambda_{ij}^{(2)} = 1, \\ 0, & \text{otherwise} \end{cases}$$

where s_j, y_j are the j-th components of vectors s, y, respectively.

Now, the factorization update algorithm is of the form

$$\begin{cases} x^{(k+1)} = x^{(k)} + s^{(k)}, \qquad s^{(k)} = x^{(k+1)} - x^{(k)}, \\ U_k s^{(k)} = -H_k F(x^{(k)}), \end{cases} \tag{1.2}$$

where, given an initial value $x^{(0)}$ and initial approximation $H_0^{-1} U_0$ to $F'(x^{(0)})$, $H_0 \in \mathcal{L}, U_0 \in \mathcal{U}$, the matrices $\{H_k\}, \{U_k\}$ are generated by

$$\begin{cases} H_{k+1} = H_k - \sum_{i=1}^{n} \langle v_k(i), v_k(i) \rangle^+ e_i^T (H_k y_k - U_k s^{(k)}) e_i y_k(i)^T, \\ U_{k+1} = U_k + \sum_{i=1}^{n} \langle v_k(i), v_k(i) \rangle^+ e_i^T (H_k y_k - U_k s^{(k)}) e_i s^{(k)}(i)^T \end{cases} \tag{1.3}$$

and

$$\begin{cases} y_k = F(x^{(k+1)}) - F(x^{(k)}), \\ v_k(i)^T = (y_k(i)_1, \cdots, y_k(i)_{i-1}, s^{(k)}(i)_i, \cdots, s^{(k)}(i)_n), \end{cases} \tag{1.4}$$

where α^+ denotes $1/\alpha$ for $\alpha \neq 0$ or 0 for $\alpha = 0$, respectively, $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product, and e_i is the i-th unit vector in R^n.

In the following discussion, we always assume that $F : D \subset R^n \to R^n$ is continuously differentiable and F' satisfies the Lipschitz condition with the Lipschitz constant γ on an open convex set $D_0 \subset D$, and there exist $H(x) \in \mathcal{L}, U(x) \in \mathcal{U}$ such that $H(x)F'(x) = U(x)$ holds for all $x \in D_0$. In addition, we denote $\|\cdot\|$ as $\|\cdot\|_2$, and $\|\cdot\|_F$ as the Frobenius