ZAMM · Z. Angew. Math. u. Mech. 64 (1984) 9, 413 – 418

ALEFELD, G.

On the Convergence of the Higher Order Versions of D. J. Evans' Implicit Matrix Inversion Process

Dedicated to Prof. Dr. Dr. h. c. HELMUT HEINRICH on the occasion of his 80th birthday.

Die Versionen des Evans-Verfahrens mit höherer Konvergenzordnung werden untersucht. Es werden hinreichende Bedingungen für die Konvergenz, Fehlerabschätzungen sowie Bedingungen für die Monotonie der Iterierten angegeben. Es wird außerdem gezeigt, daß diese Verfahren zur Verbesserung einer Näherung für die Pseudoinverse einer Matrix mit vollem Spaltenrang geeignet sind.

The higher order versions of the method introduced by D. J. Evans are investigated. We introduce sufficient conditions for convergence, prove some error estimations and show that under appropriate conditions the iterates are possessing monotonous behaviour. Furthermore we show that these methods can also be used to improve an approximation for the pseudo-inverse of a matrix with full column rank.

Изучаются варианты метода Ивеньса с высшим порядком сходимости и даются достаточные условия для сходимости, оценки погрешности и условия для монотонности итерированных. Кроме того доказывается, что эти методы пригодны для улучшения приближения для псевдо-обратной одной матрицы с полным столбцевым рангом.

1. Introduction

In the recent paper [2] D. J. EVANS introduced an implicit matrix inversion process. It was demonstrated by a simple example that this method is asymptotically much faster convergent than the wellknown method of SCHULZ [3], which is usually denoted as the Hotelling-method in the English literature. In [1] the method proposed by EVANS was investigated. Sufficient conditions for convergence, error estimations and statements about the monotonous behaviour of the iterates were proved.

In the present paper we show that EVANS' method can also be used to improve an approximation of the pseudo-inverse of an (m, n)-matrix A which has full column rank. Furthermore we investigate the versions of EVANS' method which have higher order of convergence. For these methods we present sufficient conditions for convergence and error estimations as well as statements about the monotonous behaviour of the iterates.

2. Preleminaries

Let there be given a real (m, n)-matrix A with full column rank, that is rank $(A) = n$. It follows that $m \geqq n$. The pseudoinverse of A is defined to be the unique (n, m)-matrix $A^\dagger$ for which the following equations (the so-called *Moore-Penrose equations*) hold:

$$\left.\begin{array}{ll}\text{(a)}\quad AA^\dagger A = A\,, & \text{(b)}\quad A^\dagger AA^\dagger = A^\dagger\,,\\ \text{(c)}\quad (AA^\dagger)^\mathsf{T} = AA^\dagger\,, & \text{(d)}\quad (A^\dagger A)^\mathsf{T} = A^\dagger A\,.\end{array}\right\} \qquad (1)$$

If $\operatorname{rank}(A) = n$ *and* $m = n$ *hold then* $A^\dagger = A^{-1}$. *Under the above assumption* $\operatorname{rank}(A) = n$ *the equation*

$$A^\dagger A = I_n \qquad (2)$$

holds where I_n *is the* (n, n) *unit matrix.* This can be seen in the following manner: In general we have

$$\operatorname{rank}(UV) \leqq \min\{\operatorname{rank}(U), \operatorname{rank}(V)\}$$

for the rank of the product UV. Hence it follows from (1), (a) that

$$\operatorname{rank}(A) = \operatorname{rank}(AA^\dagger A) \leqq \operatorname{rank}(A^\dagger A)$$

and therefore

$$\operatorname{rank}(A^\dagger A) \leqq \operatorname{rank}(A) \leqq \operatorname{rank}(A^\dagger A)$$

from which $\operatorname{rank}(A^\dagger A) = n$ follows. Therefore the (n, n)-matrix $A^\dagger A$ is nonsingular. Multiplying the equation

$$(A^\dagger A)(A^\dagger A) = A^\dagger AA^\dagger A = A^\dagger A$$

from the left by $(A^\dagger A)^{-1}$ the assertion follows. □

If we set

$$X = VA^\mathsf{T} \qquad (3)$$

for an (n, n)*-matrix* V *and for the given matrix* A *then it holds that*

$$X = XAA^\dagger\,. \qquad (4)$$

This can be seen in the following manner: From (1), (a) and (c) it follows that

$$A^\mathsf{T} = A^\mathsf{T}(AA^\dagger)^\mathsf{T} = A^\mathsf{T}(AA^\dagger)\,,$$

and therefore using (3),

$$X = VA^\mathsf{T}(AA^\dagger) = XAA^\dagger$$

which is (4). □

The existence of a matrix V such that (3) *holds is equivalent to the condition that the range of X^T is contained in the range of A:*

$$X = VA^\mathsf{T} \text{ for some } (n, n)\text{-matrix } V \Leftrightarrow R(X^\mathsf{T}) \subset R(A) \text{ (where } R \text{ denotes the range)}\,.$$

In order to prove this assume on the one hand that $X = VA^\mathsf{T}$. Then it follows for $x \in \mathbb{R}^n$ that $X^\mathsf{T}x = AV^\mathsf{T}x$ and because of $V^\mathsf{T}x \in \mathbb{R}^n$ it follows that $R(X^\mathsf{T}) \subset R(A)$. If on the other hand $R(X^\mathsf{T}) \subset R(A)$ then for every $\tilde{x} \in \mathbb{R}^n$ there exists an $x \in \mathbb{R}^n$ such that $X^\mathsf{T}\tilde{x} = Ax$.

Choosing $\tilde{x} = e_i \in \mathbb{R}^n$, $i = 1(1)\,n$, where e_i denotes the i-th unit-vector we get $X^\mathsf{T} = AV^\mathsf{T}$ or $X = VA^\mathsf{T}$. The i-th column of V^T is a vector x_i which because of the equation $X^\mathsf{T}e_i = Ax_i$ corresponds to $\tilde{x} = e_i$. □

If (besides of $\operatorname{rank}(A) = n$) $m = n$ then because of $R(A) = \mathbb{R}^n$ it holds for all matrices X that $R(X^\mathsf{T}) \subset R(A) = \mathbb{R}^n$, that is (3) is true for all matrices X with some V.

We now mention some results which partly where already proved in [1].

Theorem 1: *Assume that for the (n, n)-matrix $A = D - L - U$ (D diagonal part, L lower triangular part, U upper triangular part of A) the diagonal part D is nonsingular. Let $s_i \neq 0$, $i = 1(1)\,n$, and $S = \operatorname{diag}(s_i)$.*

a) *Define the real numbers p_i, $i = 1(1)\,n$, recursively by*

$$p_i = \frac{1}{|s_i|}\left\{\sum_{j=1}^{i-1}\left|\frac{a_{ij}}{a_{ii}}\right| p_j|s_j| + \sum_{j=i+1}^{n}\left|\frac{a_{ij}}{a_{ii}}\right| |s_j|\right\}, \qquad i = 1(1)\,n\,,$$

and assume that

$$p = \max_{1\le i\le n} p_i < 1\,.$$

Then using the norm $\|\cdot\| = \|S^{-1}\cdot S\|_\infty$ it holds that

$$\|(D - L)^{-1}\,U\| \leqq p < 1\,.$$

b) *Define the real numbers q_i, $i = 1(1)\,n$, recursively by*

$$q_i = \frac{1}{|s_i|}\left\{\sum_{j=1}^{i-1}\left|\frac{a_{ij}}{a_{ii}}\right| |s_j| + \sum_{j=i+1}^{n}\left|\frac{a_{ij}}{a_{ii}}\right| q_j|s_j|\right\}, \qquad i = n(-1)\,1\,,$$

and assume that

$$q = \max_{1\le i\le n} q_i < 1\,.$$

Then using the norm $\|\cdot\| = \|S^{-1}\cdot S\|_\infty$ it holds that

$$\|(D - U)^{-1}\,L\| \leqq q < 1\,.$$

Theorem 2: *Assume that the (n, n)-matrix A has nonvanishing diagonal elements and that for the numbers $s_i \neq 0$, $i = 1(1)\,n$, it holds that*

$$r = \max_{1\le i\le n}\left\{\frac{1}{|s_i|}\sum_{\substack{j=1\\ j\neq i}}^{n}\left|\frac{a_{ij}}{a_{ii}}\right| |s_j|\right\} < 1\,.$$

Then $p \leqq r < 1$ and $q \leqq r < 1$ for the numbers p and q defined in Theorem 1.

If not stated otherwise we always use the norm

$$\|\cdot\| = \|S^{-1}\cdot S\|_\infty \tag{5}$$

where S is some fixed nonsingular diagonal-matrix.

We are using the following lemmata:

Lemma 1: *Assume that A is a real (m, n)-matrix and X is a real (n, m)-matrix. If $\|I_n - XA\| < 1$ for the norm* (5) *then XA has nonvanishing diagonal elements.*

For the case $m = n$ the proof can be found in [1] as a part of the proof of Satz 1 of this paper. The analogous proof holds for $m \neq n$.

Lemma 2: *Assume that A is a real (m, n)-matrix and X is a real (n, m)-matrix. If $\|I_n - XA\| < 1$ for the norm* (5) *then*

$$\|I_n - D^{-1}XA\| = \|D^{-1}(L + U)\| < 1$$

where $XA = D - L - U$.

For the case $m = n$ the proof can be found in [1] as a part of the proof of Satz 1 of this paper. The analogous proof holds for $m \neq n$.

Lemma 3. *Assume that A is a real (m, n)-matrix and X is a real (n, m)-matrix. If $XA = D - L - U$ with nonsingular diagonal part D then if*

$$||I_n - D^{-1}XA|| = ||D^{-1}(L + U)|| < 1$$

for the norm (5) *it always holds that*

$$||I_n - D^{-1}XA|| \leqq ||I_n - XA|| \,. \tag{6}$$

Furthermore if all diagonal elements of XA are different from one then (6) *holds with the strict* $<$-sign.

Proof: By assumption $||I_n - D^{-1}XA|| = ||D^{-1}(L + U)|| < 1$, that is

$$\frac{1}{|s_i|}\left\{\sum_{j\neq i}^{n}\left(\left|\frac{l_{ij}}{d_{ii}}\right||s_j| + \left|\frac{u_{ij}}{d_{ii}}\right||s_j|\right)\right\} < 1\,, \qquad i = 1(1)\,n\,,$$

where we have set $D = \operatorname{diag}(d_{ii})$, $L = (l_{ij})$, $U = (u_{ij})$.

If $d_{ii} \neq 1$ then multiplying this inequality by $|1 - d_{ii}| \neq 0$ and using $1 - |d_{ii}| \leqq |1 - d_{ii}|$ it follows that

$$\frac{1}{|s_i|}\left\{\sum_{j\neq i}^{n}\left(\left|\frac{l_{ij}}{d_{ii}}\right||s_j| + \left|\frac{u_{ij}}{d_{ii}}\right||s_j|\right)\right\} < |1 - d_{ii}| + \frac{1}{|s_i|}\left\{\sum_{j\neq i}^{n}(|l_{ij}|\,|s_j| + |u_{ij}|\,|s_j|)\right\}.$$

If $d_{ii} \neq 1$, $i = 1(1)\,n$, then it follows that $||I_n - D^{-1}XA|| < ||I_n - XA||$. If at least one $d_{ii} = 1$ then we only can conclude that $||I_n - D^{-1}XA|| \leqq ||I_n - XA||$. □

3. Methods

Let there be given an (m, n)-matrix A with full column rank, that is rank $(A) = n \leqq m$. The (n, m)-matrix X_k is assumed to be an approximation for the pseudoinverse $A^\dagger$ of A. We are forming the product X_kA which is an (n, n)-matrix and consider the splitting

$$X_kA = D_k - L_k - U_k \tag{7}$$

where D_k is the diagonal part and L_k and U_k denote the strictly lower and upper triangular parts of X_kA, respectively. If D_k is nonsingular then we define

$$\tilde{L}_k = D_k^{-1}L_k\,, \qquad \tilde{U}_k = D_k^{-1}U_k \tag{8}$$

and

$$F_k = \tilde{L}_k\tilde{U}_k(I_n - \tilde{U}_k)^{-1}\,(I_n - \tilde{L}_k)^{-1}\,, \qquad G_k = (I_n - \tilde{L}_k)\,(I_n - \tilde{U}_k)\,. \tag{9}$$

It then holds that

$$D_k^{-1}X_kA = (I_n - F_k)\,G_k\,. \tag{10}$$

In order to improve X_k we consider for some nonnegative integer r the iteration method

$$X_{k+1} = G_k^{-1}(I_n + F_k + \dots + F_k^r)D_k^{-1}X_k\,, \qquad k = 1, 2, \dots\,. \tag{11}$$

For the special case $m = n$ and for $r = 0$ this method was proposed by EVANS in [2].

Theorem 3 (Convergence statements): *Let A be an (m, n)-matrix with full column rank. Assume that for some (n, n)-matrix V_1 it holds that*

$$X_1 = V_1A^\top \quad \text{and} \quad ||I_n - X_1A|| < 1$$

where the norm (5) *is used. Then method* (11) *is well-defined and it holds that* $\lim_{k\to\infty} X_k = A^\dagger$.

For the sequence of the residuals $I_n - X_kA$ it holds, using the norm (5), *that*

$$||I_n - X_{k+1}|| \leqq ||I_n - D_1^{-1}X_1A||^{[2(r+1)]^k} \leqq ||I_n - X_1A||^{[2(r+1)]^k}\,. \tag{12}$$

Proof: From the assumption $||I_n - X_1A|| < 1$ it follows, using Lemma 1, that X_1A has nonvanishing diagonal elements. Hence X_2 can be computed using (11). Multiplying (11) from the right by A it follows that

$$\begin{aligned} X_2A &= G_1^{-1}(I_n + F_1 + \dots + F_1^r)\,D_1^{-1}X_1A = G_1^{-1}(I_n + F_1 + \dots + F_1^r)\,(I_n - \tilde{L}_1 - \tilde{U}_1) = \\ &= G_1^{-1}(I_n + F_1 + \dots + F_1^r)\,(I_n - F_1)\,G_1 = I_n - G_1^{-1}F_1^{r+1}G_1\,, \end{aligned}$$

or

$$I_n - X_2A = G_1^{-1}F_1^{r+1}G_1 = (G_1^{-1}F_1G_1)^{r+1}\,.$$

Since the matrices $\tilde{L}_1$ and $(I - \tilde{L}_1)^{-1}$ commute we obtain using the definition of G_1 and F_1

$$I_n - X_2A = (G_1^{-1}F_1G_1)^{r+1} = [(I_n - \tilde{U}_1)^{-1}\,(I_n - \tilde{L}_1)^{-1}\,\tilde{L}_1\tilde{U}_1]^{r+1} = [(I_n - \tilde{U}_1)^{-1}\,\tilde{L}_1(I_n - \tilde{L}_1)\,\tilde{U}_1]^{r+1}\,.$$

Using Lemma 2 it follows from the assumption $||I_n - X_1A|| < 1$ that $||I_n - D_1^{-1}X_1A|| < 1$. From Theorem 2 and Lemma 3 it therefore follows that

$$||(I_n - \tilde{U}_1)^{-1}\,\tilde{L}_1|| \leqq ||I_n - D_1^{-1}X_1A|| \leqq ||I_n - X_1A||$$

and

$$||(I_n - \tilde{L}_1)^{-1} \tilde{U}_1|| \leqq ||I_n - D_1^{-1}X_1A|| \leqq ||I_n - X_1A||$$

where in both lines the second $\leqq$-sign can be replaced by the $<$-sign if all diagonalelements of X_1A are not equal to one.

Therefore we get

$$||I_n - X_2A|| \leqq ||I_n - D_1^{-1}X_1A||^{2(r+1)} \leqq ||I_n - X_1A||^{2(r+1)} < 1 .$$

Therefore for X_2 the same conditions hold as we have assumed for X_1 and X_3 can be computed.

It follows that

$$||I_n - X_3A|| \leqq ||I_n - D_2^{-1}X_2A||^{2(r+1)} \leqq ||I_n - X_2A||^{2(r+1)} ,$$

and using the preceding inequality that

$$||I_n - X_3A|| \leqq ||I_n - D_1^{-1}X_1A||^{[2(r+1)]^2} \leqq ||I_n - X_1A||^{[2(r+1)]^2} .$$

Using mathematical induction it follows for $k \geqq 1$ that

$$||I_n - X_{k+1}A|| \leqq ||I_n - D_1^{-1}X_1A||^{[2(r+1)]^k} \leqq ||I_n - X_1A||^{[2(r+1)]^k} .$$

Hence it follows that

$$\lim_{k\to\infty} (I_n - X_{k+1}A) = \lim_{k\to\infty} (G_k^{-1}F_kG_k)^{r+1} = 0 .$$

From the assumption $X_1 = V_1A^\top$ and the iteration method (11) it follows by mathematical induction that

$$X_{k+1} = V_{k+1}A^\top$$

where V_{k+1} is the (n, n)-matrix

$$V_{k+1} = G_k^{-1}(I_n + F_k + \dots + F_k^r)\, D_k^{-1}V_k .$$

Because of (3) and (4) it follows that $X_{k+1} = X_{k+1}AA^\dagger$.

Therefore, multiplying the equation

$$I_n - X_{k+1}A = (G_k^{-1}F_kG_k)^{r+1}$$

from the right by A it follows that

$$A^\dagger - X_{k+1} = (G_k^{-1}F_kG_k)^{r+1}\, A^\dagger$$

from which the assertion $\lim_{k\to\infty} X_k = A^\dagger$ follows. □

We add some remarks:

Assume that $m = n$. Then the second order *Schulz-method*

$$Y_{k+1} = Y_k + (I_n - Y_kA)\, Y_k , \qquad k = 1, 2, \dots ,$$

is convergent to A^{-1} if $\varrho(I_n - Y_1A) < 1$ where ϱ denotes the spectral radius. $\varrho(I_n - Y_1A) < 1$ holds, for example, for $Y_1 = X_1$ if $||I_n - X_1A|| < 1$. A simple computation shows that

$$||I_n - Y_{k+1}A|| \leqq ||I_n - X_1A||^{2^k} .$$

On the other hand it follows from (12) for $r = 0$ that

$$||I_n - X_{k+1}A|| \leqq ||I_n - D_1^{-1}X_1A||^{2^k} \leqq ||I_n - X_1A||^{2^k} .$$

As mentioned in the proof of Theorem 3 the second $\leqq$-sign can be replaced by the strict $<$-sign if all diagonal-elements of X_1A are different from one. This means that we have a better estimation for the residual $||I_n - X_{k+1}A||$ compared with the estimation of the residual $||I_n - Y_{k+1}A||$. *In this sense the method of* Evans *is faster than the method of* Schulz.

Theorem 4 (Monotone Convergence): *Assume that the (m, n)-matrix A has full column rank and that for some (n, m)-matrix $X_1 \geqq 0$ it holds that $I_n - X_1A \geqq 0$ where $X_1 = V_1A^\top$ for some (n, n)-matrix V_1. If $A^\dagger \geqq 0$ and if each row of X_1 contains at least one nonvanishing entry then* (11) *is well-defined and it holds that*

$$0 \leqq X_1 \leqq X_2 \leqq \dots \leqq X_k \leqq X_{k+1} \leqq \dots \leqq A^\dagger ,$$

that is $\lim_{k\to\infty} X_k = X^* \leqq A^\dagger$. If $\varrho(I_n - X_1A) < 1$ where ϱ denotes the spectral radius then $\lim_{k\to\infty} X_k = A^\dagger$.

Proof: We first show that under the assumptions $I_n - X_kA \geqq 0$, $X_k \geqq 0$, $X_k = V_kA^\top$ for some $k \geqq 0$ it holds that $0 < d_{ii} \leqq 1$ for the diagonal elements of X_kA. This can be seen as follows: From the assumption $I_n - X_kA = I_n - D_k + L_k + U_k \geqq 0$ it follows that

$$L_k \geqq 0 , \qquad U_k \geqq 0 , \qquad I_n \geqq D_k$$

and therefore that

$$X_kA = D_k - L_k - U_k \leqq D_k .$$

Since by assumption $A^\dagger \geqq 0$ it follows that

$$X_kAA^\dagger \leqq D_kA^\dagger .$$

By (3) and (4) it follows from the hypothesis $X_k = V_k A^\top$ that $X_k = X_k A A^\dagger$. Therefore the last inequality can be written as

$$X_k \leqq D_k A^\dagger .$$

Set $A^\dagger = (b_{ij})$, $X_k = (x_{ij})$ and consider a fixed row i. Then $x_{ij} \leqq d_{ii} b_{ij}$, $1 \leqq j \leqq m$, and therefore $d_{ii} \geqq 0$. If $d_{ii} = 0$ then $x_{ij} = 0$ for $1 \leqq j \leqq m$ which contradicts the assumption. Therefore X_{k+1} can be computed using (11). We now show that the assumptions $X_k \geqq 0$, $X_k = V_k A^\top$ and $I_n - X_k A \geqq 0$ imply that $X_{k+1} \geqq X_k$.

Since $D_k \geqq 0$, $L_k \geqq 0$, $U_k \geqq 0$, $X_k \geqq 0$ it follows that $F_k \geqq 0$ and therefore using (11) that

$$X_{k+1} = G_k^{-1}(I_n + F_k + \dots + F_k^r)\, D_k^{-1} X_k = (I_n - \tilde{U}_k)^{-1}\,(I_n - \tilde{L}_k)^{-1}\,(I_n + F_k + \dots + F_k^r)\, D_k^{-1} X_k =$$
$$= (I_n + \tilde{U}_k + \dots + \tilde{U}_k^{n-1})\,(I_n + \tilde{L}_k + \dots + \tilde{L}_k^{n-1})\,(I_n + F_k + \dots + F_k^r)\, D_k^{-1} X_k =$$
$$= D_k^{-1} X_k + \text{nonnegative terms} \geqq D_k^{-1} X_k \geqq X_k .$$

As in the proof of Theorem 3 one shows that

$$I_n - X_{k+1} A = (G_k^{-1} F_k G_k)^{r+1} = [(I_n - \tilde{U}_k)^{-1}\,(I_n - \tilde{L}_k)^{-1}\,\tilde{L}_k \tilde{U}_k]^{r+1}$$

which implies that $I_n - X_{k+1} A \geqq 0$ and therefore since $A^\dagger \geqq 0$

$$A^\dagger - X_{k+1} A A^\dagger \geqq 0 .$$

As in the proof of Theorem 3 it follows from $X_k = V_k A^\top$ that $X_{k+1} = V_{k+1} A^\top$ and therefore that $X_{k+1} A A^\dagger = X_{k+1}$. We therefore have $A^\dagger \geqq X_{k+1}$. Since $0 \leqq X_k \leqq X_{k+1}$ it follows that X_{k+1} has in each row at least one entry different from zero.

Therefore (11) is well defined and it holds that

$$0 \leqq X_1 \leqq X_2 \leqq \dots \leqq X_k \leqq X_{k+1} \leqq \dots \leqq A^\dagger$$

from which $\lim_{k\to\infty} X_k = X^* \leqq A^\dagger$ follows. If $\varrho(I_n - X_1 A) < 1$ then there exists a nonsingular diagonal matrix S such that $\|I_n - X_1 A\| < 1$ for the norm (5). Hence $\lim_{k\to\infty} X_k = A^\dagger$ follows from Theorem 3. □

4. Error estimations and computational amount

For simplicity we consider only the special case $m = n$ in this section. This implies that $A^\dagger = A^{-1}$. Using (10) and (11) we arrive at

$$A^{-1} - X_{k+1} - (I_n - D_k^{-1} X_k A)\,(A^{-1} - X_{k+1}) = D_k^{-1} X_k A (A^{-1} - X_{k+1}) =$$
$$= D_k^{-1} X_k - D_k^{-1} X_k A X_{k+1} = D_k^{-1} X_k - (I_n - F_k)\, G_k G_k^{-1}(I_n + F_2 + \dots + F_k^r)\, D_k^{-1} X_k =$$
$$= D_k^{-1} X_k - (I - F_k^{r+1})\, D_k^{-1} X_k = F_k^{r+1} D_k^{-1} X_k .$$

For the norm (5) the inequality $|A| \leqq |B|$ implies that $\|A\| \leqq \|B\|$. Assuming $\|I_n - D_k^{-1} X_k A\| < 1$ and using besides of this

$$|D_k^{-1} L_k| \leqq |D_k^{-1} L_k + D_k^{-1} U_k| \leqq |I_n - D_k^{-1} X_k A| ,$$
$$|D_k^{-1} U_k| \leqq |D_k^{-1} L_k + D_k^{-1} U_k| \leqq |I_n - D_k^{-1} X_k A|$$

we get

$$\|A^{-1} - X_{k+1}\| \leqq \frac{1}{1 - \|I_n - D_k^{-1} X_k A\|} \cdot \frac{\|D_k^{-1} L_k\|^{r+1}\, \|D_k^{-1} U_k\|^{r+1}}{(1 - \|D_k^{-1} L_k\|)^{r+1}\,(1 - \|D_k^{-1} U_k\|)^{r+1}}\, \|D_k^{-1} X_k\| \leqq$$
$$\leqq \frac{1}{1 - \|I_n - D_k^{-1} X_k A\|} \cdot \frac{\|I_n - D_k^{-1} X_k A\|^{2(r+1)}}{(1 - \|I_n - D_k^{-1} X_k A\|)^{2(r+1)}} \cdot \|D_k^{-1} X_k\| .$$

Since by Lemma 2 $\|I_n - X_k A\| < 1$ implies that $\|I_n - D_k^{-1} X_k A\| < 1$ we have the following result.

Theorem 5 (Error estimation): *Assume that for some (n, n)-matrix A there exists an (n, n)-matrix X_1 such that for the norm (5) $\|I_n - X_1 A\| < 1$. Then for the iterates computed by (11) it follows that*

$$\|A^{-1} - X_{k+1}\| \leqq \frac{1}{1 - \|I_n - D_k^{-1} X_k A\|} \cdot \frac{\|D_k^{-1} L_k\|^{r+1} \|D_k^{-1} U_k\|^{r+1}}{(1 - \|D_k^{-1} L_k\|)^{r+1}\,(1 - \|D_k^{-1} U_k\|)^{r+1}}\, \|D_k^{-1} X_k\| \leqq$$
$$\leqq \frac{\|I_n - D_k^{-1} X_k A\|^{2(r+1)}}{(1 - \|I_n - D_k^{-1} X_k A\|)^{2(r+1)+1}}\, \|D_k^{-1} X_k\| \leqq \frac{\|I_n - X_k A\|^{2(r+1)}}{(1 - \|I_n - X_k A\|)^{2(r+1)+1}} \cdot \|D_k^{-1} X_k\| .$$

Proof: For the proof we only have to note that $\|I_n - X_1 A\| < 1$ implies $\|I_n - X_k A\| < 1$ for all $k \geqq 1$. This was already proved as a part of the proof of Theorem 3. □

From Theorem 5 it follows that the order of convergence of (11) is (at least) $2(r + 1)$.

We are now discussing the computational amount necessary to perform one step of the method (11). For simplicity we again assume that $m = n$. As it is general use we only count the so-called point operations (multiplications and divisions) and neglect lower order powers of n.

1. The computation of X_kA needs n^3 operations.
2. Computation of F_k: Writing the equation (9) as

$$F_k(I_n - D_k^{-1}L_k)\,(I_n - D_k^{-1}U_k) = D_k^{-1}L_kD_k^{-1}U_k$$

we get F_k by first solving the equation

$$S_k(I_n - D_k^{-1}U_k) = D_k^{-1}L_kD_k^{-1}U_k$$

for S_k and subsequently solving the equation

$$F_k(I_n - D_k^{-1}L_k) = S_k$$

for F_k. Hence we have to solve two matrix equations with triangular matrices as coefficient matrices. This needs altogether n^3 operations if the right hand sides are known. For getting the right hand sides we have to compute the product $D_k^{-1}L_kD_k^{-1}U_k$ of the lower triangular matrix $D_k^{-1}L_k$ and the upper triangular matrix $D_k^{-1}U_k$. For this one has to perform $\frac{1}{3}\,n^3$ operations. Therefore the computation of F_k needs $n^3 + \frac{1}{3}\,n^3 = \frac{4}{3}\,n^3$ operations.

3. Computation of $I_n + F_k + \dots + F_k^r$ using HORNER's scheme needs $r - 1$ matrix multiplications that is $n^3(r - 1)$ operations.
4. Computation of $(I_n + F_k + \dots + F_k^r)\,D_k^{-1}X_k$ needs one matrix multiplication that is n^3 operations.
5. Computation of X_{k+1}: Writing (11) as

$$G_kX_{k+1} = (I_n + F_k + \dots + F_k^r)\,D_k^{-1}X_k =: R_k\,,$$

then X_{k+1} can be computed by solving

$$(I_n - D_k^{-1}L_k)\,Z_k = R_k$$

for Z_k and subsequently solving

$$(I_n - D_k^{-1}U_k)\,X_{k+1} = Z_k\,.$$

This needs n^3 operations.

Adding the operations from steps 1. to 5. we have the result that for $r > 0$

$$n^3 + \tfrac{4}{3}\,n^3 + n^3(r - 1) + n^3 + n^3 = n^3\,(\tfrac{10}{3} + r)$$

operations are necessary to perform one step of (11).

Note that for $r = 0$ steps 2), 3) and 4) are not performed. Therefore one step of (11) needs $2n^3$ operations if $r = 0$.

We now compare the amount of work with that necessary to perform one step of the higher order version of the SCHULZ method: If r is a nonnegative integer then the method

$$Y_{k+1} = [I_n + (I_n - Y_kA) + \dots + (I_n - Y_kA)^{r+1}]\,Y_k\,, \qquad k = 1,2,\dots\,, \tag{13}$$

is convergent if $\varrho(I_n - Y_1A) < 1$. The order of convergence is $r + 2$.

The computational amount of work necessary to perform one step of this method is as follows:

1. Computation of Y_kA needs n^3 operations.
2. The term [...] can be computed by using HORNER's scheme. This needs $r \cdot n^3$ operations.
3. Multiplying the term [...] by Y_k needs n^3 operations.

Therefore altogether one step of (13) needs

$$n^3 + rn^3 + n^3 = n^3(r + 2)$$

operations. This result also holds for $r = 0$.

In the following Table we list for some values of r the amount of work necessary to perform one step of (11) and (13). Furthermore the Table contains the order of convergence for these values of r.

r	0	1	2	3	4	5	6
(13)	$2n^3$	$3n^3$	$4n^3$	$5n^3$	$6n^3$	$7n^3$	$8n^3$
Order of Convergence	2	3	4	5	6	7	8
(11)	$2n^3$	$\frac{13}{3}\,n^3$	$\frac{16}{3}n^3$	$\frac{19}{3}n^3$	$\frac{22}{3}\,n^3$	$\frac{25}{3}\,n^3$	$\frac{28}{3}\,n^3$
Order of Convergence	2	4	6	8	10	12	14

References

1 ALEFELD, G., Zur Konvergenz eines Verfahrens von D. J. EVANS zur iterativen Verbesserung einer Näherung für die inverse Matrix, Numer. Math. **39** (1982) 163–173.
2 EVANS, D. J., An implicit matrix inversion process, SIGNUM Newsletter **15** (1980) 3.
3 SCHULZ, G., Iterative Berechnung der reziproken Matrix, ZAMM **13** (1933) 57–59.

Received October 10, 1983

Address: Prof. Dr. G. ALEFELD, Institut für Angew. Mathematik, Universität Karlsruhe, Kaiserstr. 12, D-7500 Karlsruhe, F.R.G.