Appl. Math. Lett. Vol. 6, No. 4, pp. 83–86, 1993
0893-9659/93 $6.00 + 0.00

A NOTE ON ITERATIVE REFINEMENT SCHEMES FOR SYLVESTER OPERATOR EQUATIONS

MARIO AHUES AND ALAIN LARGILLIER
Université Jean Monnet, Saint-Etienne, France
BALMOHAN LIMAYE
Indian Institut of Technology, Bombay, India

(Received and accepted December 1992)

Abstract—We propose two iterative schemes to refine an approximate solution of a Sylvester operator equation $\underline{K}x - x\theta = y$, where K is a bounded linear operator in a Banach space B, $\underline{K}$ its extension to the product space $X = B^m$, and $\theta \in \mathbb{C}^{m\times m}$. An approximate solution x_n is obtained by means of an approximation K_n to K. Then, x_n is refined by two iterative processes involving the resolution, for e_n, of $\underline{K_n}e_n - e_n\theta = r_n$, with different second members r_n. In these processes, K is only used for evaluations.

1. MATHEMATICAL BACKGROUND

Let $(B, |\cdot|)$ be a Banach space over the complex field $\mathbb{C}$, $(\mathcal{L}(B), |\cdot|)$, the space of bounded linear operators in B, m a positive integer, and $X = B^m$ the product space. For $x = (x_1, \dots, x_m) \in X$, we set $\|x\| = (\sum_{i=1}^m |x_i|^2)^{1/2}$, and given an operator $A : B \to B$, we write $\underline{A}$ its natural extension to X : $\underline{A}x = (Ax_1, \dots, Ax_m)$. It can be shown that $A \in \mathcal{L}(B)$ implies $\underline{A} \in \mathcal{L}(X)$ and $\|\underline{A}\| = |A|$. For $A \in \mathcal{L}(B)$, $\mathrm{sp}(A)$ is the spectrum of A and $\mathrm{re}(A)$ the resolvent set. For $z \in \mathrm{re}(A)$, $R(A, z) = (A - zI)^{-1} \in \mathcal{L}(B)$ is the resolvent. For $\theta = (\theta_{ij}) \in \mathbb{C}^{m\times m}$, we define $x\theta = (\sum_{i=1}^m \theta_{i1}x_i, \dots, \sum_{i=1}^m \theta_{im}x_i) \in X$, and the following inequality holds: $\|x\theta\| \le \|x\||\theta|_F$, where $|\cdot|_F$ is the Frobenius norm. Let there be given $K \in \mathcal{L}(B)$ and $\theta \in \mathbb{C}^{m\times m}$ such that θ is invertible and

$$\mathrm{sp}(K) \cap \mathrm{sp}(\theta) = \emptyset. \tag{1.1}$$

We define the linear operator $G : X \to X$ by $Gx = \underline{K}x - x\theta$. Hypothesis (1.1) implies that G has an inverse $G^{-1} \in \mathcal{L}(X)$. We are interested in solving the equation

$$Gx = y, \tag{1.2}$$

where $y \in X$ is given. Let $K_n \in \mathcal{L}(B)$, $n \in \mathbb{N}$, be a sequence of operators. We define $G_n \in \mathcal{L}(X)$ by

$$\forall x \in X, \qquad G_n x = \underline{K_n}x - x\theta. \tag{1.3}$$

In what follows, we suppose that

$$K \in \mathcal{L}(B) \text{ is a compact operator}, \tag{1.4}$$
$$K_n \text{ is pointwise convergent to } K, \tag{1.5}$$
$$(K_n - K)K_n \text{ converges in norm to } 0. \tag{1.6}$$

We shall prove that under these hypotheses, given any $y \in X$, the approximate equation

$$G_n x_n = y \tag{1.7}$$

is uniquely solvable for all n large enough.

Typeset by $\mathcal{A}_\mathcal{M}\mathcal{S}$-TEX

THEOREM 1.1. *For each $z \in \mathrm{re}(K)$ there exists $n(z) \in \mathbb{N}$ such that for all $n > n(z)$, $z \in \mathrm{re}(K_n)$ and $c(z) = \sup_{n>n(z)} |R(K_n, z)|$ is a finite constant.*

PROOF. See [1]. ∎

THEOREM 1.2. *There exists $n(\theta) \in \mathbb{N}$ such that $C_\theta = \sup_{n>n(\theta)} \|G_n^{-1}\|$ is a finite constant.*

PROOF. By Schur's theorem, there exists a unitary matrix $Q \in \mathbb{C}^{m\times m}$ such that $\tau = Q^*\theta Q = (\tau_{ij})$ is upper triangular. Condition (1.1) implies that for n large enough and $j = 1, \dots, m$, $\tau_{jj} \in \mathrm{re}(K_n)$. The equation (1.7) can be written as

$$\underline{K_n} x'_n - x'_n \tau = y', \tag{1.8}$$

where $x'_n = x_n Q$ and $y' = yQ$. If $x' = (x'_{1n}, \dots, x'_{mn})$ and $y' = (y'_1, \dots, y'_m)$, then the solution $x'_n \in X$ is given by

$$x'_{1n} = R(K_n, \tau_{11})\, y'_1, \quad x'_{jn} = R(K_n, \tau_{jj}) \left(y'_j + \sum_{i=1}^{j-1} \tau_{ij} x'_{in} \right), \qquad j = 2, \dots, m.$$

Since $x = x'Q^*$ and $|Q^*|_F = |Q|_F = \sqrt{m}$, the result follows. ∎

THEOREM 1.3. *x_n converges to x as n tends to infinity.*

PROOF. Since $x_n - x = (G_n^{-1} - G^{-1})\, y = G_n^{-1}(G - G_n)\, G^{-1} y = G_n^{-1}(\underline{K} - \underline{K_n})x$, and since $\underline{K_n}$ is pointwise convergent to $\underline{K}$, then the uniform boundedness of G_n^{-1} is sufficient to the convergence of x_n to x. ∎

2. ITERATIVE REFINEMENT SCHEMES

THEOREM 2.1. *For n large enough but fixed, the iterative refinement scheme*

$$x^{(0)} = x_n = G_n^{-1} y, \quad x^{(k+1)} = x^{(k)} - G_n^{-1}(\underline{K} x^{(k)} - x^{(k)}\theta - y), \qquad k \ge 0, \tag{2.1}$$

converges linearly to x as $k \to \infty$. Moreover, there exist $\alpha > 0$ and $\gamma_n \in\,]0, 1[$ such that, for all $k \ge 0$, $\max\{\|x^{(2k)} - x\|, \|x^{(2k+1)} - x\|\} \le \alpha(\gamma_n)^k$.

PROOF. We have $x^{(k+2)} - x = \left[G_n^{-1}(\underline{K_n} - \underline{K})\right]^2 (x^{(k)} - x)$. But $\left[G_n^{-1}(\underline{K_n} - \underline{K})\right]^2 = G_n^{-1}((\underline{K_n} - \underline{K})\underline{K_n} G_n^{-1} + (\underline{K} - \underline{K_n}) G_n^{-1} \underline{K})$, which tends to 0 in the norm of $\mathcal{L}(X)$ since G_n^{-1} is uniformly bounded, $(K_n - K)K_n$ tends to 0 in the norm of $\mathcal{L}(B)$, $(\underline{K} - \underline{K_n})G_n^{-1}$ is pointwise convergent to 0, and $\underline{K}$ is compact because K is compact. ∎

THEOREM 2.2. *For n large enough but fixed, the iterative refinement scheme*

$$\begin{aligned} x^{(0)} &= x_n = G_n^{-1} y, \quad x^{(k+1/2)} = (\underline{K} x^{(k)} - y)\, \theta^{-1}, \\ x^{(k+1)} &= x^{(k+1/2)} - G_n^{-1}(\underline{K} x^{(k+1/2)} - x^{(k+1/2)}\theta - y), \qquad k \ge 0, \end{aligned} \tag{2.2}$$

converges linearly to x as $k \to \infty$. Moreover, there exist $\beta > 0$ and $\delta_n \in\,]0, 1[$ such that, for all $k \ge 0$, $\|x^{(k)} - x\| \le \beta(\delta_n)^k$.

PROOF. We have $x^{(k+1)} - x = G_n^{-1}\left((\underline{K_n} - \underline{K})\underline{K}\right)\left((x^{(k)} - x)\theta^{-1}\right)$. But $K_n - K$ is pointwise convergent to 0, K is compact, and G_n^{-1} is uniformly bounded so that $G_n^{-1}(\underline{K_n} - \underline{K})\underline{K}$ converges to 0 in the norm of $\mathcal{L}(X)$. ∎

3. NUMERICAL EXAMPLES

Let $(\omega_{jn})_{j=1}^n$ be the weights and $(t_{jn})_{j=1}^n$ the knots of a quadrature formula, pointwise convergent on the space $\mathcal{C}[0, 1]$ of complex valued continuous functions defined on $[0, 1]$. Let K be an integral operator defined by a continuous kernel κ and K_n the Nyström approximation associated with the given quadrature formula:

$$\forall \varphi \in \mathcal{C}[0,1], \qquad (K_n\varphi)(s) = \sum_{j=1}^{n} \omega_{jn}\kappa(s,t_{jn})\,\varphi(t_{jn}), \qquad s \in [0,1].$$

All the hypotheses of Section 1 are then satisfied. Computational experiments have been done with the compounded trapezoidal quadrature rule based upon n equally spaced knots. The kernel κ and the matrix θ are given by

$$\kappa(s,t) = \begin{cases} 10t(1-s), & \text{if } 0 \le t \le s \le 1, \\ 10s(1-t), & \text{if } 0 \le s < t \le 1, \end{cases} \quad \text{and} \quad \theta = \begin{pmatrix} \lambda & \nu & 0 \\ 0 & \lambda & 0 \\ \nu & \nu & \lambda \end{pmatrix}.$$

Hence, $\operatorname{sp}(\theta) = \{\lambda\}$, and the departure from normality is of the order of ν^2. The second member $y = (y_1, y_2, y_3)$ is given by $y_1(s) = \sin 10s$, $y_2(s) = e^s$, and $y_3(s) = s^2$. Iterations have been stopped when the residual is less than $5.0 \cdot 10^{-14}$. Evaluations of K have been done with a fine discretization K_N, with $N \gg n$. Table 3.1 shows the number of iterations performed by each method for different values of n and N. Table 3.2 shows the first twelve residuals and their ratios for each method in one of the cases in Table 3.1.

Table 3.1. Number of iterations needed to obtain a residual less than $5.0E-14$.

n	N	λ	ν	Method A	Method B
3	100	−1.0	0.0	30	16
3	100	−1.0	10.0	40	21
5	100	−1.0	10.0	22	11
10	150	1.0	0.0	78	36
10	200	0.8	4.0	34	16
10	250	2.0	20.0	15	7

Table 3.2. Residuals and their ratios in the case $n = 5$, $N = 100$.

Iteration	Residual of A	Ratio	Residual of B	Ratio
0	$7.9E+0$		$5.7E+0$	
1	$3.0E+0$	0.38	$5.0E-1$	0.08
2	$7.5E-1$	0.25	$3.6E-2$	0.07
3	$2.4E-1$	0.32	$2.1E-3$	0.05
4	$4.9E-2$	0.20	$1.1E-4$	0.05
5	$1.3E-2$	0.26	$5.4E-6$	0.05
6	$2.5E-3$	0.19	$2.6E-7$	0.05
7	$6.2E-4$	0.25	$1.2E-8$	0.05
8	$1.1E-4$	0.17	$5.2E-10$	0.04
9	$2.6E-5$	0.24	$2.2E-11$	0.04
10	$4.7E-6$	0.18	$9.5E-13$	0.04
11	$1.0E-6$	0.21	$3.7E-14$	0.04
12	$1.8E-7$	0.18		

4. FINAL COMMENTS

Method (2.2) has a better rate of convergence than (2.1). This phenomenon was observed in the case of Fredholm equations of the second kind [2]. However, (2.2) needs one additional evaluation of K. Method (2.2) appears to be more stable than (2.1). There exist ill-conditioned situations in which (2.1) diverges and (2.2) converges. We recall that the condition number of (1.7) depends on the departure from normality of θ. In conclusion, we suggest that (2.2) should be prefered to (2.1). For the numerical resolution of (1.7), the reader is refered to [3].

REFERENCES

1. M. Ahues, A class of strongly stable approximations, *J. Austral. Math. Soc. Ser. B* **28**, 435–442 (1987).
2. H. Brakhage, Über die numerische Behandlung von Integralgleichungen nach der Quadraturformelmethode, *Numer. Math.* **2**, 183–196 (1960).
3. G. Golub, S. Nash and C. Van Loan, A Hessenberg-Schur method for the problem $AX + XB = C$, *IEEE Trans. Autom. Control* **AC-24**, 909–913 (1979).