FRIEDRICH-ALEXANDER-UNIVERSITÄT ERLANGEN-NÜRNBERG
INSTITUT FÜR INFORMATIK (MATHEMATISCHE MASCHINEN UND DATENVERARBEITUNG)

Lehrstuhl für Informatik 10 (Systemsimulation)

An Iterative Algorithm for Approximate Orthogonalisation of Symmetric Matrices

M. Mohr, C. Popa, and U. Rüde

Lehrstuhlbericht 03-2

Abstract

In a previous paper one of the authors presented an extension of an iterative approximate orthogonalisation algorithm, due to Z. Kovarik, for arbitrary rectangular matrices. In the present paper we propose a modified version of this extension, for the class of arbitrary symmetric matrices. For this new algorithm, the computational effort per iteration is much smaller than for the initial one. We prove its convergence and also derive an error reduction factor per iteration. In the second part of the paper we show that we can eliminate the matrix inversion required by the previous algorithm in each iteration, by replacing it with a polynomial matrix expression. Some numerical experiments are also presented for a collocation discretisation of a first kind integral equation.

AMS Subject Classification: 65F10, 65F20

Key words: orthogonalisation algorithm, symmetric matrix, approximate inverse

1 Extended Kovarik algorithm

We start this paper by introducing some notation and shortly presenting the extension, see [5], of Kovarik's original algorithm B from [3]. Let A be an $m \times n$ matrix, then we shall denote by $(A)_i$, A^t, and $A^\dagger$ its i-th row, transpose and Moore-Penrose pseudo-inverse, see [1], respectively. We define as generalised spectral condition number $k_2(A)$ the square root of the ratio between the largest and smallest singular value of A, and $\langle \cdot, \cdot \rangle$, $\| \cdot \|$ will be the Euclidean scalar product and norm on some space $I\!R^q$. For a square matrix B, $\sigma(B)$ will denote its spectrum and $\rho(B)$ its spectral radius. All vectors appearing in the paper are understood to be column vectors. With this preparation, we can now give the extension of Kovarik's algorithm B, that was proposed in [5].

Algorithm 1 (KOB) *Let $A_0 = A$ be a general matrix. Then for $k = 0, 1, \ldots,$ we construct a sequence via*

$$K_k = (I - A_k A_k^t)(I + A_k A_k^t)^{-1}, \quad A_{k+1} = (I + K_k) A_k. \tag{1}$$

Let us now suppose that the matrix A satisfies

$$\| AA^t \|_2 = \rho(AA^t) < 1, \tag{2}$$

where $\| AA^t \|_2$ denotes its spectral norm. Then, using its smallest nonzero eigenvalue $\lambda_{\min}(AA^t)$, we can define

$$\delta = 1 - \lambda_{\min}(AA^t).$$

and $\delta > 0$ will hold. The following result was proved in [5].

Theorem 1 *If A satisfies (2), then the sequence $(A_k)_{k \geqslant 0}$ constructed in (1) converges and*

$$A_\infty := \lim_{k \to \infty} A_k = [(AA^t)^{\frac{1}{2}}]^+ A \ . \tag{3}$$

Moreover, the following estimate holds

$$\| A_k - A_\infty \|_2 \leqslant \delta^{2^k} \quad , \quad \forall k \geqslant 1 \ . \tag{4}$$

Remark 1 *Relation (4) shows, that the convergence of the KOB algorithm is quadratic. Moreover, the assumption (2) is not restrictive. It can be obtained by a scaling of the matrix A of the form*

$$A^{new} := \frac{1}{\sqrt{\| A \|_\infty \| A \|_1 + 1}} A \ ,$$

where $\| \cdot \|_\infty$ and $\| \cdot \|_1$ are the well known matrix norms, derived from their vector counterparts, see e.g. [1].

Assume, that $U^t A V = \mathrm{diag}(\sigma_1, \ldots, \sigma_r, 0, \ldots, 0)$ with $r = \mathrm{rank}(A)$ is a singular value decomposition of A and $\tilde{I}$ is the $m \times m$ matrix defined by

$$\tilde{I} = \mathrm{diag}(1, 1, \ldots, 1, 0, \ldots, 0) \tag{5}$$

then, it was proven in [5], that the following "approximate orthogonalisation" relation holds with respect to the rows of the matrix A_∞ from (3)

$$\langle (A_\infty)_i, (A_\infty)_j \rangle = \langle \tilde{I}(U)_i, (U)_j \rangle \ , \tag{6}$$

which for $\tilde{I} = I$, i.e. for A with linearly independent rows, becomes a classical orthogonality, since the matrix U is orthonormal. Moreover, the following result can also be proved with respect to the generalised spectral condition number of the matrices A_k from (1)

$$\lim_{k\to\infty} k_2(A_k) = k_2(A_\infty) = 1. \tag{7}$$

2 Modified Kovarik algorithm for symmetric matrices

Let us now suppose that A is an $n \times n$ symmetric matrix. Then, if A satisfies (2) the above algorithm KOB can be applied to it and all the results from the previous section hold. Moreover, it can easily be proved that all the matrices A_k will be symmetric. Thus (1) will become

$$K_k = (I - A_k^2)(I + A_k^2)^{-1}, \quad A_{k+1} = (I + K_k)A_k. \tag{8}$$

However, computing in every step the product A_k^2 represents a large computational effort. We therefore now present the following modified version of the KOB algorithm (8), which eliminates this need.

Algorithm 2 (KOBS) *Let $A_0 = A$ be a general symmetric matrix. Then for $k = 0, 1, \ldots,$ we construct a sequence via*

$$K_k = (I - A_k)(I + A_k)^{-1} \ , \quad A_{k+1} = (I + K_k)A_k \ . \tag{9}$$

Before considering the well-definedness and convergence properties of the algorithm KOBS, we start with the following auxiliary result.

Lemma 1 *The real sequence $x_{k+1} := \Phi(x_k)$ with $k \geqslant 0$, $x_0 \in I\!R$ and defined by the function*

$$\Phi(x) := \frac{2x}{1+x}$$

has the following properties:

1. *There are two fixed points 0 and 1, of which 0 is repellent and 1 attractive.*
2. *For any initial value x_0 in the set*
$$E := \left\{ -\frac{1}{\alpha_j}, j \in \mathbb{N}_0 \ \middle| \ \alpha_0 = 1, \alpha_{j+1} = 2\alpha_j + 1 \right\} \tag{10}$$
the sequence will break down at the point -1, since $\Phi(-1)$ is not well-defined. Furthermore there exists no $v \in \mathbb{R} \setminus E$ such that $\Phi(v) \in E$.
3. *For any initial guess $x_0 \in \mathbb{R} \setminus E$ the sequence will converge to $\lim_{k\to\infty} x_k = 1$.*
4. *For every starting value $x_0 \in (-1, 0) \setminus E$ there exists an index k^* such that $x_{k^*} < -1$ and $x_{k^*+1} > 1$.*

Proof: It is obvious to see, that $\Phi(x) = x$ holds for $x = 0$ and $x = 1$, and only for these values. So 0 and 1 are fixed points. In order to see their behaviour, we distinguish the following four cases.

(i) Let $1 < x_k < \infty$. In this situation, we get

$$x_{k+1} < x_k \Longleftrightarrow \frac{2x_k}{1+x_k} < x_k \Longleftrightarrow \frac{2}{1+x_k} < 1 \Longleftrightarrow 2 < 1 + x_k \Longleftrightarrow 1 < x_k$$

and

$$1 < x_{k+1} \Longleftrightarrow 1 < \frac{2x_k}{1+x_k} \Longleftrightarrow 1 + x_k < 2x_k \Longleftrightarrow 1 < x_k$$

so in this situation, the sequence will converge towards 1 from above.

(ii) Let $0 < x_k < 1$. In this situation we have

$$x_{k+1} > x_k \Longleftrightarrow \frac{2x_k}{1+x_k} > x_k \Longleftrightarrow \frac{2}{1+x_k} > 1 \Longleftrightarrow 2 > 1 + x_k \Longleftrightarrow x_k < 1$$

and

$$x_{k+1} < 1 \Longleftrightarrow \frac{2x_k}{1+x_k} < 1 \Longleftrightarrow 2x_k > 1 + x_k \Longleftrightarrow x_k < 1$$

so in this situation, the sequence will converge towards 1 from below.

(iii) Let $-\infty < x_k < -1$. Now in this case, the next iterate will be larger than 1, since,

$$x_{k+1} > 1 \Longleftrightarrow \frac{2x_k}{1+x_k} > 1 \Longleftrightarrow 2x_k < 1 + x_k \Longleftrightarrow x_k < 1$$

so for the next iterate x_{k+1} we can apply (i) to see, that the sequence will converge to 1.

(iv) Let $-1 < x_k < 0$. In this case, we see from

$$x_{k+1} < x_k \Longleftrightarrow \frac{2x_k}{1+x_k} < x_k \Longleftrightarrow \frac{2}{1+x_k} > 1 \Longleftrightarrow 2 > 1 + x_k \Longleftrightarrow x_k < 1$$

that there will be a number $n \in \mathbb{N}$, such that $x_{k+n} \in (-\infty, -1]$. For convergence however, we need $x_{k+n} \neq -1$, since $\Phi(-1)$ is obviously not well-defined. A quick calculation shows, that $\Phi(x) = -1$ exactly for $x = -1/3$. Recursion of this computation leads to the set E for which the following holds

$$\Phi(-\frac{1}{\alpha_{j+1}}) = -\frac{1}{\alpha_j} \quad .$$

So for any $x_0 \in E$ the sequence (x_k) will finally lead us to -1 and break down. Since the function Φ is strictly monotonically increasing ($\Phi'(x) = 2/(1+x)^2$) there can be no other value $v \in (-1,0) \setminus E$ such that $\Phi(v) \in E$. So for all $x_0 \in (-1,0) \setminus E$ the sequence will decrease until for some index k^* we get $x_{k^*} < -1$.

Combining the results of the four cases completes the proof.

□

Using the above lemma, it will be can now toggle the question of convergence of the algorithm KOBS.

Theorem 2 *Let us suppose that the matrix A is symmetric and that none of its eigenvalues lies in the set E from (10). Then, the sequence $(A_k)_{k\geqslant 0}$ generated by the algorithm (9) it well-defined and it converges towards*

$$A_\infty := \lim_{k\to\infty} A_k = A^\dagger A.$$

Proof: Let Q be an orthonormal matrix such that

$$Q^t A Q = Q^t A_0 Q = D_0 = \operatorname{diag}(\lambda_1^{(0)}, \ldots, \lambda_r^{(0)}, 0, \ldots, 0) \ , \tag{11}$$

where $\lambda_i^{(0)}$ are the nonzero eigenvalues of A and $\operatorname{rank}(A) = r$. Assuming, that the sequence exists up to an index k, we obtain by recursion, that

$$A_k = Q D_k Q^t, \ D_k = \operatorname{diag}(\lambda_1^{(k)}, \ldots, \lambda_r^{(k)}, 0, \ldots, 0) \ . \tag{12}$$

Moreover the values $\lambda_i^{(k)}$, $i = 1, \ldots, r$ are given by

$$\lambda_i^{(k)} = \frac{2\lambda_i^{(k-1)}}{1 + \lambda_i^{(k-1)}} = \Phi\left(\lambda_i^{(k-1)}\right) \ , \tag{13}$$

with the function Φ from Lemma 1. The next element A_{k+1} of the sequence (9) will be well-define, when the matrix $(I + A_k)$ is invertible. The latter will be the case, when none of the eigenvalues of

A_k is equal to -1. Thus property 2 of Lemma 1 shows, that for any symmetric matrix A the KOBS algorithm will be well-defined, as long as no eigenvalue of A lies in the set E.

Concerning the convergence of the sequence A_k, we see from (13), that this is a direct consequence of property 3 of Lemma 1. So, if A fulfils the requirements of the theorem, then

$$A_\infty := \lim_{k\to\infty} A_k$$

exists and

$$Q^t A_\infty Q = \operatorname{diag}(\underbrace{1,\dots,1}_{r \text{ times}},0,\dots,0)$$

However, we obtain from (11)

$$Q^t A^\dagger Q = \operatorname{diag}\left(\frac{1}{\lambda_1},\dots,\frac{1}{\lambda_r},0,\dots,0\right) \quad ,$$

and thus

$$A^\dagger A = Q\operatorname{diag}(1,\dots,1,0,\dots.0)\,Q^t = A_\infty \tag{14}$$

This completes the proof.

□

The convergence of the algorithm KOBS is no longer quadratic as for KOB in (4). Instead we have the following result.

Corollary 1 *Let A be a symmetric matrix with $\sigma(A)\cap E=\emptyset$. If also $\sigma(A)\cap(0,1)=\emptyset$ holds, then if $(A_k)_{k\geqslant 0}$ is the sequence generated by the KOBS algorithm we have*

$$\| A_k - A^\dagger A \|_2 \quad \leqslant \quad \left(\frac{1}{2}\right)^k \| A_0 - A^\dagger A\|_2 \ , \quad \forall k \geqslant k_0 \ , \tag{15}$$

where $k_0=0$ if no eigenvalue of A is negative and $k_0=k^*+2$ with the k^* from Lemma 1 otherwise. If $\sigma(A)\subset[0,1]$ we instead get the following. Let $\delta>0$ be given by

$$\delta = \frac{1}{1+\lambda_{min}(A)}, \tag{16}$$

where $\lambda_{min}(A)\in(0,1)$ is the smallest positive eigenvalue of A. Then,

$$\| A_k - A^\dagger A \|_2 \quad \leqslant \quad \delta^k \ , \quad \forall k\geqslant 0 \ . \tag{17}$$

Proof: We firstly observe that $A^\dagger A$ from (14) can be written as

$$A^\dagger A = Q\tilde{I}Q^t,$$

with $\tilde{I}$ from (5). Then, also using (12) we get

$$\| A_k - A^\dagger A \|_2 \quad = \quad \| D_k - \tilde{I} \|_2 \quad = \quad \max_{1\leqslant i\leqslant r} |\lambda_i^{(k)} - 1| \ , \tag{18}$$

with $r=\operatorname{rank}(A)$. In order to evaluate $|\lambda_i^{(k)}-1|$ in (18) we come back to Lemma 1 and its proof. We first consider the case that $\sigma(A)\cap(0,1)=\emptyset$. Assume that $x_0=\lambda_i^{(0)}>1$. In this case, the sequence $(x_k)_k$ is strictly monotonically decreasing to 1 and thus

$$|x_{k+1}-1| = \frac{x_k-1}{1+x_k} < \frac{1}{2}\,|x_k-1| \ , \quad \forall k\geqslant 0$$

and by a recursive argument we obtain

$$|x_{k+1}-1| \leqslant \left(\frac{1}{2}\right)^{k+1} |x_0-1| \ , \quad \forall k\geqslant 0 \ . \tag{19}$$

If $x_0 = \lambda_i^{(0)} < 0$, then we know from Lemma 1 that there exists an index k^* such that $x_{k^*+1} > 1$. So (19) will hold for all $k > k^* + 1$. Let us now consider the second case $\sigma(A) \subset (0, 1)$. Here we have $x_0 = \lambda_i^{(0)} \in (0, 1)$ for $i = 1, \dots, r$ and δ from (16) is well-defined. From the proof of Lemma 1 we know, that in $(0, 1)$ the sequence $(x_k)_k$ is strictly monotonically increasing. Thus we get

$$|x_k - 1| = \left| \frac{2x_{k-1}}{1 + x_{k-1}} - 1 \right| = \frac{|x_{k-1} - 1|}{1 + x_{k-1}} < \frac{1}{1 + \lambda_{\min}(A)} |x_{k-1} - 1| \ .$$

So by a recursive argument we get

$$|x_k - 1| < \left(\frac{1}{1 + \lambda_{\min}(A)} \right)^k |x_0 - 1| \ , \quad \forall k \geqslant 0 \ .$$

Since $|x_0 - 1| < 1$ the proof is complete.

□

Remark 2 *If $\sigma(A) \cap (0, 1) \neq \emptyset$ and also $\sigma(A) \cap (I\!R \setminus (0, 1)) \neq \emptyset$ we have to combine the two estimates (15) and (17).*

Remark 3 *The following "approximate orthogonalisation" relation, see also (6), holds for the matrix $A^\dagger A$*

$$\langle (A^\dagger A)_i, (A^\dagger A)_j \rangle = \langle \tilde{I}(Q)_i, (Q)_j \rangle \ ,$$

with Q from (11) and $\tilde{I}$ from (5). Moreover, we have a similar result as in (7) with respect to the generalised spectral condition number of A_k, namely

$$\lim_{k \to \infty} k_2(A_k) = k_2(A^\dagger A) = 1.$$

Remark 4 *It is well known, see e.g. [1], that*

$$A^\dagger A = P_{R(A^t)},$$

where $P_{R(A^t)}$ is the orthogonal projection onto the subspace $R(A^t) \subset I\!R^n$, i.e. the range of A^t. But in our case $A = A^t$, thus

$$A^\dagger A = P_{R(A)} \ , \quad I - A^\dagger A = P_{N(A)} \ .$$

From Theorem 2 it then holds that, for a given vector $b \in I\!R^n$ we have

$$\lim_{k \to \infty} A_k b = P_{R(A)}(b) \ , \quad \lim_{k \to \infty} (b - A_k b) = P_{N(A)}(b) \ .$$

Remark 5 *Assume that $\sigma(A) \subset (0, 1)$. Corollary 1 shows, that in this case, the convergence of $\lambda_i^{(k)}$ to one will be the slower, the closer $\lambda_i^{(0)}$ is to zero. Now in the situation that A is positive (semi)-definite and has both eigenvalues smaller and larger than one, the latter will converge much more quickly than the former. This aspect might be of interest in so called discrete ill-posed problems coming from the discretisation of ill-posed inverse problems, since it constitutes an implicit regularisation of a sort. See [2] in this respect.*

3 Modified Kovarik algorithm without matrix inversion

Although the algorithm KOBS is less costly from the point of computational effort than KOB, it still contains the unpleasant step related to the inversion of the matrix $(I + A_k)$ in each iteration. The modified version of KOBS, which we shall present in this section, will eliminate this part. In this section we restrict ourselves to positive semi-definite matrices, which satisfy

$$\sigma(A) \subset [0, 1] \ . \tag{20}$$

For such matrices, we know that all matrices A_k created by the KOBS algorithm will also satisfy (20). Furthermore, it is a well known fact, see e.g. [1], that for a matrix B with $\rho(B) < 1$ the inverse of $(I + B)$ can be expressed via the von Neumann series

$$(I+B)^{-1} = \sum_{j=0}^{\infty} (-B)^j \ . \tag{21}$$

The idea now is the following. Given a fixed sequence $(n_k)_{k\geqslant 0}$ of positive integers, we approximate the infinite series in (21) by $\sum(A_n; n_k)$ defined as

$$\sum(A_n; n_k) = \sum_{j=0}^{n_k} (-A_k)^j \ , \quad \forall k \geqslant 0 \ .$$

This leads us to the following modified version of the algorithm KOBS.

Algorithm 3 (MKOBS) *Let $A_0 = A$ be a symmetric matrix, which satisfies (20). Then for $k = 0, 1, \dots,$ we construct a sequence via*

$$K_k = (I - A_k) \sum(A_k; n_k) \ , \quad A_{k+1} = (I + K_k) A_k \ . \tag{22}$$

Before we start analysing the convergence properties of the MKOBS algorithm, we give an auxiliary result.

Lemma 2 *Let $(n_k)_{k\geqslant 0}$ be a sequence of positive integers. Using these numbers we define two other sequences $\mathcal{S}_1 := (y_k)_{k\geqslant 0}$ and $\mathcal{S}_2 := (z_k)_{k\geqslant 0}$ by*

$$y_{k+1} := \frac{2y_k - y_k^{n_k+2}(y_k - 1)}{y_k + 1} \tag{23}$$

$$z_{k+1} := \frac{2z_k + z_k^{n_k+2}(z_k - 1)}{z_k + 1} \ . \tag{24}$$

Let us denote by x_k an element of $\mathcal{S}_1$ or $\mathcal{S}_2$ respectively. Now both sequences $\mathcal{S}_1$ and $\mathcal{S}_2$ have the following properties:

1. *If 0 and 1 is chosen as starting value of the sequence, all elements of the sequence will be 0 or 1 respectively.*
2. *For any initial value x_0 with $0 < x_0 < 1$ the sequence x_{k+1} will converge strictly monotonically towards 1 from below.*

Proof: It is easy to see that $x_0 = 0$ leads to $x_k = 0$ for all $k > 0$ and that this also holds for $x_0 = 1$. We now assume that $0 < x_k < 1$. We start by proving that in this case both sequences are bounded by 1 from above. This can easily be seen from

$$\begin{aligned} x_{k+1} < 1 \quad &\Longleftrightarrow \quad \frac{2x_k \pm x_k^{n_k+2}(x_k - 1)}{x_k + 1} < 1 \\ &\Longleftrightarrow \quad 2x_k \pm x_k^{n_k+2}(x_k - 1) < x_k + 1 \\ &\Longleftrightarrow \quad \pm x_k^{n_k+2}(x_k - 1) < 1 - x_k \\ &\Longleftrightarrow \quad \mp x_k^{n_k+2}(1 - x_k) < 1 - x_k \\ &\Longleftrightarrow \quad \pm x_k^{n_k+2} < 1 \end{aligned}$$

by applying recursion. Next we show, that both sequences increase strictly monotonically for $0 < x_0 < 1$. Let us start with $\mathcal{S}_1$. In this case we have

$$y_{k+1} - y_k = \frac{y_k(1 - y_k)(1 + {y_k}^{n_k+1})}{1 + y_k} > \frac{y_k(1 - y_k)}{1 + y_k} \ . \tag{25}$$

From the boundedness of y_k we see by recursion, that the right hand side of this equation is positive, thus the first sequence must grow strictly monotonically. In the case of $\mathcal{S}_2$ we get

$$z_{k+1} - z_k = \frac{z_k(1-z_k)(1-{z_k}^{n_k+1})}{1+z_k} > \frac{z_k(1-z_k)^2}{1+z_k} \quad , \tag{26}$$

which by the same argument shows, that also the second sequence is strictly monotonically growing. So both sequences will converge and it remains to be shown, that their limits y^* and z^* are equal to 1. Let us instead assume that they satisfy

$$0 < y^* < 1 \ , \quad 0 < z^* < 1 \ . \tag{27}$$

In this case we get from (25) and (26)

$$\begin{aligned} y_{k+1} - y_k &\geqslant \frac{y_k(1-y_k)}{1+y^*} \\ y_{k+1} - y_k &\geqslant \frac{z_k(1-z_k)^2}{1+z^*} \ . \end{aligned}$$

Studying the behaviour of the functions $f(t) = t(1-t)$ and $g(t) = t(1-t)^2$ for $t \in (0,1)$ we can conclude, that

$$\begin{aligned} y_{k+1} - y_k &\geqslant \frac{\tau_y}{1+y^*} \\ z_{k+1} - z_k &\geqslant \frac{\tau_z}{1+z^*} \ , \end{aligned}$$

with $\tau_y := \min\{f(y_0),\, f(y^*)\} > 0$ and $\tau_z := \min\{g(z_0),\, g(z^*)\} > 0$. The latter holds due to $y_0 > 0$, $z_0 > 0$, and (27). Thus, we see that none of the sequences is a Cauchy sequence, which is a contradiction to them being convergent. Thus the assumption is wrong and we have $y^* = 1 = z^*$.

□

Theorem 3 *Let A be symmetric and positive semi-definite, such that (20) holds. Assume furthermore, that the sequence $(n_k)_{k\geqslant 0}$ consists only of either odd or even positive integers. Then, the sequence $(A_k)_{k\geqslant 0}$ generated by (22) converges to $A^\dagger A$.*

Proof: Let Q be an orthonormal matrix consisting of eigenvectors of A, then (11) holds with

$$\lambda_i^{(0)} \in (0,1] \ , \quad \forall i = 1,2,\ldots,r \ ,$$

with $r = \operatorname{rank}(A)$. As in the proof of Theorem 1 we obtain that A_{k+1} is of the form (12) with D_{k+1} given by

$$D_{k+1} = \left(h^{(k)}(\lambda_1^{(k)}),\ldots,h^{(k)}(\lambda_r^{(k)}),0,\ldots,0\right) \ ,$$

where for MKOBS $h^{(k)}$ is given by

$$h^{(k)}(x) = \begin{cases} \dfrac{2x + x^{n_k+2}(x-1)}{x+1} \ , & \text{if } n_k \text{ is odd} \\[2ex] \dfrac{2x - x^{n_k+2}(x-1)}{x+1} \ , & \text{if } n_k \text{ is even} \ . \end{cases} \tag{28}$$

Thus the convergence behaviour of the MKOBS algorithms is determined by that of the sequences $\mathcal{S}_1$ and $\mathcal{S}_2$ from Lemma 2. This shows, that for matrices that satisfy (20) the method will converge.

□

In the case that the sequence (n_k) consists only of even integers we can derive a convergence estimate for the MKOBS algorithm that is similar to that for KOBS.

Corollary 2 *Let the sequence $(n_k)_{k\geqslant 0}$ of positive integers consists only of even numbers, and let A be a symmetric matrix, that satisfies (20). Then with δ defined by (16) we have*

$$\| A_k - A^\dagger A \|_2 \ \leqslant \ \delta^k \ , \quad \forall k \geqslant 0 \ . \tag{29}$$

Proof: The proof essentially follows the proof of Corollary 1. We again use the orthonormal matrix Q of eigenvectors of A to arrive at the representation (18) of $\| A_k - A^\dagger A \|_2$. In order to evaluate the development of the norm of this difference, we have to consider the sequence $\mathcal{S}_1$ of Lemma 2, since all n_k are even, cf. the proof of Theorem 3. We get that for any $i = 1, \ldots, r$ and $\lambda_i = y_0$ the following holds

$$|\lambda_i^{(k)} - 1| = 1 - y_k = \frac{1 - y_{k-1}}{1 + y_{k-1}}(1 - y_{k-1}^{n_k+2}) < \frac{1 - y_{k-1}}{1 + y_{k-1}} < \frac{1 - y_{k-1}}{1 + y_0} < \frac{1 - y_{k-1}}{1 + \lambda_{\min}} \ ,$$

which by a recursive argument finally yields $|\lambda_i^{(k)} - 1| < \delta^k$.

□

Remark 6 *Corollary 2 shows, that in the case of even n_k's the convergence of the MKOBS algorithm is as least as fast as that of the KOBS algorithm. Remember, however, that the former does not require the computation of an inverse matrix in every iteration step.*

Remark 7 *In the case that the sequence (n_k) consists only of odd integers we no longer have a similar theoretical convergence behaviour for the MKOBS algorithm as for KOBS. This is because for $(z_k)_k$ from (24) we get*

$$1 - z_{k+1} = \frac{1 + z_k^{n_k+2}}{1 + z_k}\,(1 - z_k)$$

and unfortunately $\lim_{k\to\infty} \frac{1+z_k^{n_k+2}}{1+z_k} = 1$.

4 Numerical experiments

We consider in our tests the following first kind integral equation. For a given function $y \in L^2([0,1])$ find $x^* \in L^2([0,1])$ such that

$$\int_0^1 k(s,t)x(t)dt = y(s) \ , \quad s \in [0,1] \ , \tag{30}$$

with

$$k(s,t) = \frac{1}{1 + |s-t|} \ , \quad y(s) = ln[(1+s)(2-s)] \ . \tag{31}$$

Remark 8 *Equation (30) has the solution $x(t) = 1$ for all $t \in [0,1]$.*

We discretise (30)-(31) by a collocation algorithm, see e.g. [4], with the collocation points

$$s_i = \frac{i-1}{n-1} \ , \quad i = 1, 2, \ldots, n \ .$$

Thus we obtain the symmetric positive definite system

$$Ax = b,$$

with the $n \times n$ matrix A and $b \in \mathbb{R}^n$ given by

$$A_{ij} = \int_0^1 k(s_i,t)k(s_j,t)dt \ , \quad b_i = y(s_i) \ .$$

The generalised condition number $k_2(A)$ is presented in Table 1 for different values of n. We now test the algorithms KOB, in the form (8), KOBS (9) and MKOBS (22), using the following three stopping rules

$$\| A_{k+1} - A_k \|_\infty \leqslant 10^{-6} \ , \tag{32}$$
$$k_2(A) \leqslant 10 \ , \tag{33}$$
$$k_2(A) \leqslant 100 \ . \tag{34}$$

For the algorithm MKOBS we use

$$n_k = Q = \text{const} \ , \quad \forall k \geqslant 0 \ ,$$

with the two different values $Q = 3$ and $Q = 5$. For MKOBS we scale the matrix A according to Remark 1 in order to obtain (20). Tables 2, 3 and 4 present the number of iterations necessary to fulfil the stopping rules (32) - (34).

Note. All the computations were made with the Numerical Linear Algebra software package OCTAVE, freely available under the terms of the GNU General Public License, see `www.octave.org`.

Table 1. The condition number of A

n	$k_2(A)$
16	$3.7\ 10^5$
32	$6.7\ 10^6$
64	$1.14\ 10^8$
128	$1.8\ 10^9$
256	$3.0\ 10^{10}$

Table 2. Results for alg. KOB (8)

n	stop (32)	stop (33)	stop (34)
16	20	12	9
32	23	16	12
64	26	19	15
128	29	22	18

Table 3. Results for alg. KOBS (9)

n	stop (32)	stop (33)	stop (34)
16	39	16	13
32	43	20	17
64	47	24	21
128	51	28	25

Table 4. Results for alg. MKOBS (22)

n	stop (32)	stop (32)	stop (33)	stop (33)	stop (34)	stop (34)
	Q=3	Q=5	Q=3	Q=5	Q=3	Q=5
16	739	609	16	16	13	13
32	748	618	20	20	17	17
64	757	626	24	24	21	21
128	763	632	28	28	25	25

5 Acknowledgements

The paper was supported by NATO through collaborative linkage grant PST.CLG.977924 and by the DAAD via a grant that one of the authors had as a visiting professor at the Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany, in the period October 2002-August 2003.

References

[1] Golub, G. H. and van Loan, C. F., *Matrix computations*, The John's Hopkins Univ. Press, Baltimore, 1983.

[2] Hansen, P. C., *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*, SIAM Monographs on Mathematical Modeling and Computation, SIAM, 1998

[3] Kovarik, S., *Some iterative methods for improving orthogonality*, SIAM J. Num. Anal., **7**(3)(1970), 386-389.

[4] Nashed, M. Z. and Wahba, G,, *Convergence rates of approximate least squares solutions of linear integral and operator equations of the first kind*, Math. of Comput., **28**(125)(1974), 69-80.

[5] Popa, C., *A method for improving orthogonality of rows and columns of matrices*, Intern. J. Comp. Math., **77**(3)(2001), 469-480.