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

Lehrstuhl für Informatik 10 (Systemsimulation)

A Kovarik Type-Algorithm without Matrix Inversion for the Numerical Solution of Symmetric Least-Squares Problems

M. Mohr, C. Popa and U. Rüde

Lehrstuhlbericht 05-2

Marcus Mohr* Constantin Popa† Ulrich Rüde, ‡

April 11, 2005

Abstract

In a previous paper we presented a version of Kovarik's approximate orthogonalization algorithm for arbitrary symmetric matrices. We also proposed a modification of this algorithm that eliminates the requirement to perform in each iteration step an explicit matrix inversion. In the present report we propose another version of this modified algorithm. This variant has the advantage of a smaller bound on the convergence factor, while the computational costs per iteration are even less than in the initial modification.

In the second part of the report we investigate the application of the new algorithm for the numerical solution of linear least-squares problems with a symmetric matrix. The basic idea is to modify also the right hand side of the problem during the transformation of the matrix. We prove that the sequence of vectors generated in this way converges to the minimal norm solution of the problem.

AMS Subject Classification: 65F10, 65F20

Key words and phrases: linear least-squares problem, minimal norm solution, orthogonalization algorithm, symmetric matrix

1 Kovarik algorithm for symmetric matrices

We begin our exposition by introducing some notation. Let A be an $n \times n$ symmetric matrix and $(A)_i$, $A^\dagger$ its i-th row and Moore-Penrose pseudoinverse (see e.g. [Bjö96]), respectively. By $gk_2(A)$ we shall denote the generalized spectral condition number of A defined as the square root of the ratio between the largest and smallest of its singular value. $\langle \cdot, \cdot \rangle, \| \cdot \|$ will be the Euclidean scalar product and norm on some space $\mathbb{R}^q$ and P_S will denote the orthogonal projection onto a vector subspace $S \subset \mathbb{R}^q$. For a square matrix B, $N(B)$, $R(B)$, $\sigma(B)$, $\rho(B)$ and $\|B\|$ represent its null space, range, spectrum, spectral radius and spectral norm, respectively.

In our previous paper [MPR04] we considered the following version of Kovarik's algorithm (B) from [Kov70].

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

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

The following results concerning the convergence properties of the above algorithm were proved in [MPR04].

Theorem 1 *Let us suppose that none of the eigenvalues of the symmetric matrix A lies in the set E of real numbers defined by*

$$E := \left\{ -\frac{1}{\alpha_j}, j \in \mathbb{N}_0 \;\middle|\; \alpha_0 = 1, \alpha_{j+1} = 2\alpha_j + 1 \right\}$$

Then the sequence of matrices $(A_k)_{k \geqslant 0}$ generated by the above algorithm KOBS converges and

$$\lim_{k \to \infty} A_k = A^\dagger A \ . \tag{1}$$

*University of Erlangen-Nuremberg, Germany, *Marcus.Mohr@lse.eei.uni-erlangen.de*
†Ovidius University, Constanţa, Romania, *cpopa@univ-ovidius.ro*
‡University of Erlangen-Nuremberg, Germany, *Ulrich.Ruede@informatik.uni-erlangen.de*

Corollary 1 *Let the same assumptions hold as in the above theorem then the following are true.*

1. *If the spectrum of A satisfies*
$$\sigma(A) \cap (0,1) = \emptyset \ ,$$
then there exists an integer $k_0 \geqslant 0$ such that
$$\|A_k - A^\dagger A\| \ \leqslant \ \left(\frac{1}{2}\right)^{k-k_0} \|A_{k_0} - A^\dagger A\|, \ \forall k \geqslant k_0 \ .$$

2. *In the case that*
$$\sigma(A) \subset [0,1]$$
we have
$$\|A_k - A^\dagger A\| \ \leqslant \ \delta^k, \ \forall k \geqslant 0 \ .$$
with the factor δ given by
$$\delta := \frac{1}{1+\lambda_{\min}(A)} > 0 \ , \tag{2}$$
and $\lambda_{\min}(A) \in (0,1)$ being the smallest positive eigenvalue of A.

Remark 1 *The above corollary tells us about the linear convergence rate of the algorithm KOBS. Moreover, if A is positive semi-definite and has eigenvalues both smaller and larger than 1, the latter will converge much quicker than the former. This aspect might be of interest in so called* discrete ill-posed problems *coming from the discretization of ill-posed inverse problems, since it constitutes an implicit regularization of a sort. See e.g. [Han98] in this respect.*

Remark 2 *For the sequence of matrices $(A_k)_{k\geqslant 0}$ generated by the algorithm KOBS it can be proved that*
$$\lim_{k\to\infty} gk_2(A_k) = gk_2(A^\dagger A) = 1 \ .$$
Thus it acts as an iterative preconditioner for A. From this point of view one can combine it with some other direct or iterative solver for a problem in which A is the system matrix. See [EP01] in this respect.

For all the good properties of the KOBS algorithm, mentioned above, there is one aspect that significantly reduces its attractivity. This is of course the requirement to perform an inversion of the matrix $I + A_k$ in each step of the iteration.

For the special case of $\sigma(A) \subset [0,1]$ we proposed in [MPR04] a modified version of KOBS which in we denoted by MKOBS. In the following we will briefly describe this modified approach. We start from the observation that, according to the proof of Theorem 1, if $\sigma(A) \subset [0,1]$ holds for the spectrum of A then
$$\|A_k\|_2 = \rho(A_k) < 1, \ \forall k \geqslant 0 \ .$$
Thus, $I + A_k$ is invertible and its inverse can be represented by the following Neumann series, see e.g. [GvL96],
$$(I + A_k)^{-1} = \sum_{j=0}^{\infty} (-A_k)^j \ .$$
Let now $(n_k)_{k\geqslant 0}$ be a fixed sequence of positive integers and $\sum(A_n; n_k)$ defined by
$$\sum(A_n; n_k) = \sum_{j=0}^{n_k} (-A_k)^j, \ \forall k \geqslant 0 \ . \tag{3}$$

Then the modified version of the algorithm KOBS is given by

Algorithm 2 (MKOBS) *Let $A_0 = A$ be a symmetric matrix, which satisfies $\sigma(A) \subset [0,1]$. Then for $k = 0, 1, \ldots,$ we construct a sequence via*
$$K_k = (I - A_k) \sum(A_k; n_k) \ , \quad A_{k+1} = (I + K_k)A_k \ . \tag{4}$$

The following convergence result for the MKOBS algorithm was proved in [MPR04].

Theorem 2 *Let A be symmetric and positive semi-definite such that $\sigma(A) \subset [0,1]$ holds and suppose that all the numbers n_k are either all even or all odd. Then, the sequence $(A_k)_{k\geqslant 0}$ generated by (13) converges to $A^\dagger A$.*

Corollary 2 *Let the assumptions of Theorem 2 be fulfilled, but let $(n_k)_{k\geqslant 0}$ be a sequence of only even numbers. Then with δ defined as in Corollary 1 we have*

$$\|A_k - A^\dagger A\|_2 \leqslant \delta^k \ , \quad \forall k \geqslant 0 \ . \tag{5}$$

Remark 3 *The above corollary demonstrate that the convergence speed of the MKOBS algorithm can be as fast as that of the original KOBS algorithm for the case of even numbers n_k. However, for odd numbers there need no longer be a linear convergence and for arbitrary integers n_k the sequence $(A_k)_{k\geqslant 0}$ from (4) may even diverge, see again [MPR04] for details.*

2 An alternate inversion-free variant

Let us suppose that the matrix A fulfills the assumptions of Theorem 2. The idea for the alternate version of an inversion-free KOBS algorithm is to replace the Neumann series expansion $\sum(A_n; n_k)$ of the inverse (3) used in MKOBS by the expression

$$S(A_k; q) := \sum_{i=0}^{q} a_i(-A_k)^i \ . \tag{6}$$

The latter stems from the original paper by Kovarik [Kov70], in which the following Taylor's expansion

$$\frac{1}{\sqrt{1-x}} = a_0 + a_1 x + \dots \tag{7}$$

with

$$a_0 = 1,\ a_{j+1} = \frac{2j+1}{2j+2} a_j, j \geqslant 0 \tag{8}$$

is used, which holds for all $x \in (-1,1)$. Employing the above replacement, the new inverse-free version of the MKOBS algorithm (4) reads as follows.

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

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

With the same technique as for the KOBS algorithm from [MPR04], we reduce the analysis of the convergence behavior of the IFKOBS algorithm to the study of the convergence properties of the sequence of real numbers given by

$$x_0 \in (0,1) \ , \quad x_{k+1} = \left[1 + (1 - x_k) \sum_{i=0}^{n_k} a_i(-x_k)^i \right] x_k,\ k \geqslant 0 \ , \tag{9}$$

with parameters a_k from (8). We begin with the following

Lemma 1 *For any $x_0 \in (0,1)$ and any $n_k \geqslant 1$ the sequence (9) converges and*

$$\lim_{k\to\infty} x_k = 1.$$

Proof: Let $k \geqslant 0$ be arbitrary but fixed and suppose that $0 < x_k < 1$. First off all we note that, using (6), we can re-write (7) for $x \in (-1,1)$ as

$$\frac{1}{\sqrt{1+x}} = \sum_{i\geqslant 0} a_i(-x)^i = S(-x; \infty) \ . \tag{10}$$

Taking into account that $S(-x;\infty)$ is an alternating series it is easy to see that for any odd number n_k we have

$$0 < 1 - \frac{1}{2}x_k = S(x_k, 1) \leqslant S(x_k; n_k) < S(x_k; \infty) = \frac{1}{\sqrt{1+x_k}} \ , \tag{11}$$

since $0 < x_k < 1$. From this we immediately obtain

$$1 - x_k S(x_k; n_k) > 1 - \frac{x_k}{\sqrt{1+x_k}} = \frac{1 + x_k - x_k^2}{\sqrt{1+x_k}(\sqrt{1+x_k} + x_k)} > 0 \ .$$

This together with the fact that subtracting one from both sides in (9) leads to

$$x_{k+1} - 1 = (x_k - 1)[1 - x_k S(x_k; n_k)] \ , \tag{12}$$

now tells us that x_{k+1} as well as x_k satisfies

$$0 < x_{k+1} < 1 \ . \tag{13}$$

If on the other hand n_k is an even number we get that

$$0 < \frac{1}{\sqrt{1+x}} = S(x_k; \infty) < S(x_k; n_k) \leqslant S(x_k; 0) = 1 \ . \tag{14}$$

Thus

$$0 < 1 - x_k < 1 - x_k S(x_k; n_k),$$

which together with (12) gives us again (13). By mathematical induction we then obtain that

$$x_k \in (0, 1), \ \forall k \geqslant 0 \ . \tag{15}$$

Let us now consider the difference between to members of the sequence (9). The latter is given by

$$x_{k+1} - x_k = x_k(1 - x_k) S(x_k; n_k) \ . \tag{16}$$

As we have seen from (11) and (14) the partial sum $S(x_k; n_k)$ is positive for any n_k. Thus, we obtain from (16) that

$$0 < x_k < x_{k+1}, \forall k \geqslant 0 \ . \tag{17}$$

This proves that the sequence $(x_k)_{k\geqslant 0}$ is monotonically increasing. From (16) it is obvious that it must converge to a limit $x^* \in (0, 1]$. We will now show that $x^* = 1$ by a contradiction argument. Let us suppose that

$$0 < x^* < 1. \tag{18}$$

We again exploit the fact that $S(x_k; \infty)$ is an alternating series to derive from (16) that the difference between two members of our $(x_k)_{k\geqslant 0}$ is bounded from below by

$$x_{k+1} - x_k \geqslant x_k(1 - x_k)\left(1 - \frac{x_k}{2}\right) \tag{19}$$

for all $k \leqslant 0$. An simple analysis of the function $x \mapsto g(x) := x(1-x)(1-x/2)$ for $x \in (0, 1)$ gives us that

$$g(x_k) \geqslant \min\{g(x_0), g(x^*)\} =: \varepsilon$$

holds for all $k \geqslant 0$. From the fact that x_0 and x^* are assumed to lie in $(0, 1)$ we have

$$g(x_k) \geqslant \varepsilon > 0 \ . \tag{20}$$

Combining (19) and (20) by a recursive argument gives us

$$x_{k+1} \geqslant x_0 + (k+1)\varepsilon, \quad \forall k \geqslant 0 \ .$$

This, however, is a clear contradiction to the boundedness of the sequence $(x_k)_{k\geqslant 0}$, since it would lead to $x_k \to \infty$ for $k \to \infty$. Thus, the assumption (18) is false, $x^* = 1$ and the proof is complete. □

Lemma 2 *For any $x_0 \in (0,1)$ and any $n_k \geqslant 1$ the sequence (9) satisfies the following*

$$|x_{k+1}-1| \leqslant |x_k - 1| \cdot \begin{cases} \left(1 - x_0 + \frac{1}{2}x_0^2\right) & \text{for odd } n_k \\ \left(1 - \frac{x_0}{\sqrt{1+x_0}}\right) & \text{for even } n_k \end{cases} .$$

Proof: The proof is a direct consequence of the proof of Lemma 1. From (12) we obtain

$$|x_{k+1}-1| = |x_k - 1| \cdot |1 - x_k S(x_k; n_k)| \ . \tag{21}$$

Let us now first consider the case that n_k is an odd number. In this situation we can use (11) to bound the second term on the right of (21) from above by

$$|1 - x_k S(x_k; n_k)| \leqslant \left|1 - x_k + \frac{1}{2}x_k^2\right| \leqslant \left(1 - x_0 + \frac{1}{2}x_0^2\right) \ .$$

Here the latter inequality can be derived from an analysis of the function $h_1(x) = x + x^2/2$ on the interval $[0,1]$. If, on the other side n_k is an even number, we can use (14) and, considering the function $h_2(x) = 1 - x/\sqrt{1+x}$, instead obtain

$$|1 - x_k S(x_k; n_k)| \leqslant \left|1 - \frac{x_k}{\sqrt{1+x_k}}\right| \leqslant \left(1 - \frac{x_0}{\sqrt{1+x_0}}\right) \ .$$

□

Theorem 3 *Let A be symmetric and positive semi-definite such that $\sigma(A) \subset [0,1]$ holds. Then the sequence of matrices $(A_k)_{k\geqslant 0}$ generated by the algorithm IFKOBS converges to $A^\dagger A$. Moreover, the convergence is linear, i.e.*

$$\|A_k - A^\dagger A\| \ \leqslant \gamma^k \|A - A^\dagger A\|, \quad \forall k \geqslant 0 \ , \tag{22}$$

with

$$\gamma = \max\left\{1 - \lambda_{\min}(A) + \frac{1}{2}\lambda_{\min}(A)^2 \ , \ 1 - \frac{\lambda_{\min}(A)}{\sqrt{1+\lambda_{\min}(A)}}\right\} \ . \tag{23}$$

Proof: Let $k \geqslant 0$ be arbitrary but fixed. Since A is symmetric, there exists an orthonormal matrix Q such that

$$Q^t A Q = Q^t A_0 Q = D_0 = \operatorname{diag}(\lambda_1^{(0)}, \dots, \lambda_r^{(0)}, 0, \dots, 0) \ , \tag{24}$$

where $\lambda_i^{(0)} \in (0,1)$ are the nonzero eigenvalues of A and $\operatorname{rank}(A) = r$. Assuming, that the sequence exists up to the index k, we obtain by recursion, that

$$A_k = Q D_k Q^t, \quad D_k = \operatorname{diag}(\lambda_1^{(k)}, \dots, \lambda_r^{(k)}, 0, \dots, 0) \ . \tag{25}$$

From the definition of the IFKOBS algorithm one easily derives that the values $\lambda_i^{(k)}$, $i = 1, \dots, r$ are given by

$$\lambda_i^{(k+1)} = \left[1 + \left(1 - \lambda_i^{(k)}\right) S\left(\lambda_i^{(k)}; n_k\right)\right] \lambda_i^{(k)} \ . \tag{26}$$

Thus, using Lemma 1 with $x = \lambda_i^{(k)}$ we obtain

$$\lim_{k\to\infty} A_k = \lim_{k\to\infty} Q D_k Q^t = Q \left(\lim_{k\to\infty} D_k\right) Q^t = Q \Delta Q^t \ ,$$

with $\Delta := \operatorname{diag}(1, \dots, 1, 0, \dots, 0)$. Since $Q\Delta Q^t = A^\dagger A$ this proves the convergence. Let us now consider the speed of the convergence. Employing the representation (25) we get

$$\|A_{k+1} - A^\dagger A\| = \|Q(D_{k+1} - \Delta)Q^t\| = \|D_{k+1} - \Delta\| = \max_{1\leqslant i\leqslant r} \left|\lambda_i^{(k+1)} - 1\right| \ .$$

Using Lemma 2 we first obtain

$$\max_{1\leqslant i\leqslant r}\left|\lambda_i^{(k+1)}-1\right|\leqslant\gamma\max_{1\leqslant i\leqslant r}\left|\lambda_i^{(k)}-1\right|=\gamma\,\|A_k-A^\dagger A\|$$

and then by a recursive argument (22) - (23).

□

Corollary 3 *For a symmetric matrix $A=A_0$ with $\sigma(A)\in[0,1]$ the matrices A_k generated by the IFKOBS algorithm are also symmetric with $\sigma(A)\in[0,1]$. Moreover we have $N(A)=N(A_k)$ for the null-space of the matrices.*

Proof: The corollary is a direct consequence of the representation of A_k in the form (25) and the convergence properties of the eigenvalues $\lambda_i^{(k)}$.

□

Remark 4 *Comparing the new algorithm IFKOBS to the first inverse-free version MKOBS we can state the following:*

1. *While for the MKOBS algorithm a linear convergence is only given for sequences of even numbers $(n_k)_{k\geqslant 0}$, IFKOBS shows a linear convergence for any choice $n_k>0$.*
2. *In order to minimize the computational work per step we want to choose n_k as small as possible, while retaining a linear convergence speed. For IFKOBS this minimal value is $n_k=1$ which leads to*
$$K_k=(I-A_k)(I-1/2\cdot A_k),\quad A_{k+1}=(I+K_k)A_k\ . \tag{27}$$
We denote this algorithm by IFKOBS(1). In the case of MKOBS on the other hand we need even n_k's and thus the minimal value is $n_k=2$ which gives
$$\begin{aligned}K_k&=(I-A_k)\left(I-A_k+A_k^2\right)\\&=(I-A_k)[\,I-A_k(I+A_k)\,]\\A_{k+1}&=(I+K_k)A_k\end{aligned}\tag{28}$$
and will be denoted by MKOBS(2). In both cases the dominating amount of work is related to the incurred matrix-matrix multiplications. For a symmetric matrix $A\in\mathbb{R}^{n\times n}$ the costs for the latter are given by $n^3+(n-1)n/2$. Comparing the two update formulas we see that IFKOBS(1) requires two matrix-matrix multiplications while MKOBS(2) requires three. Thus, the costs per step behave like $2n^3+\mathcal{O}(n^2)$ for IFKOBS(1) versus $3n^3+\mathcal{O}(n^2)$ for MKOBS(2).
3. *In the case of linear convergence the bound on the convergence factor is given by*
$$\delta=\frac{1}{1+\lambda_{\min}(A)}$$
in the case of MKOBS(2), see Corollary 2, and
$$\tilde\gamma=1-\frac{\lambda_{\min}(A)}{\sqrt{1+\lambda_{\min}(A)}}$$
for IFKOBS(1) as can be derived from Theorem 3. Comparing both factors it is easy to see that $\gamma<\delta$, which yields a smaller bound and potentially faster convergence for IFKOBS(1) than for MKOBS(2) and even KOBS, see Corollary 1.

3 Numerical solution of least-squares problems

Let A be as in Theorem 3 and $b\in\mathbb{R}^n$ a given vector. We consider the least-squares problem: find $x^*\in\mathbb{R}^n$ such that

$$\|Ax^*-b\|=\min! \tag{29}$$

Let $LSS(A;b)$ be the set of all its least-squares solutions and x_{LS} the corresponding unique minimal norm solution. If (29) is consistent we shall denote $LSS(A;b)$ by $S(A;b)$. In order to apply the above algorithm IFKOBS for the numerical solution of (29) we shall consider the following formulation of it, in which also the right hand side b is modified, see also [MPR03b].

Algorithm 4 (Algorithm IFKOBS-rhs) *Let $A_0 = A$, $b^0 = b$. For $k = 0, 1, \ldots$ do*

$$K_k = (I - A_k)S(A_k; n_k),\ A_{k+1} = (I + K_k)A_k,\ b^{k+1} = (I + K_k)b^{(k)}\ . \tag{30}$$

Now we are able to prove two results related to the approximation of x_{LS} from (29).

Theorem 4 *If the problem (29) is consistent, i.e. $b \in R(A)$, then the sequence $(b^{(k)})_{k\geqslant 0}$ from (30) converges and*

$$\lim_{k\to\infty} b^{(k)} = A^{\dagger}b = x_{LS}. \tag{31}$$

Proof: It follows from the consistency of (29) that

$$b = Ax, \tag{32}$$

for some $x \in \mathbb{R}^n$. Then, by also using (30) we obtain

$$b^1 = (I + K_0)\,b^0 = (I + K_0)A_0x = A_1x$$

and, by an induction argument

$$b^{(k)} = A_kx,\ \forall k \geqslant 0. \tag{33}$$

Employing Theorem 3 we now get

$$\lim_{k\to\infty} b^{(k)} = \left(\lim_{k\to\infty} A_k\right) x = A^{\dagger}Ax = A^{\dagger}b = x_{LS}\ .$$

□

This was the consistent case. Before turning to the inconsistent case we make the following preparations. Consider the family of functions $G_k(x)$, $k \geqslant 0$, which implicitly appear in (9), and is given by

$$G_k(x) = 1 + (1 - x)\sum_{i=0}^{n_k} a_i(-x)^i\ . \tag{34}$$

We observe that for each k the function G_k can continuously be extended to $x = 0$ via

$$G_k(0) = 2\ . \tag{35}$$

Now we can formulate the following

Theorem 5 *In the inconsistent case for (29) we have*

$$\lim_{k\to\infty} A_kb^{(k)} = x_{LS}. \tag{36}$$

Proof: Because A is symmetric we have $N(A^t) = N(A)$. Thus,

$$b = P_{R(A)}(b) + P_{N(A)}(b)\ , \tag{37}$$

with

$$P_{R(A)}(b) = Ax\ , \tag{38}$$

for some $x \in \mathbb{R}^n$. As in the proof of Theorem 3, let us denote by $r \leqslant n$ the rank of A and let Q be an orthonormal $n \times n$ matrix such that

$$A = Q\ \mathrm{diag}\left(\lambda_1^{(0)}, \ldots, \lambda_r^{(0)}, 0, \ldots, 0\right) Q^t\ . \tag{39}$$

Here $\lambda_i^{(0)}$ are the nonzero eigenvalues of A. From the representation of A in the form (25) and (30) it is easy to see that we can write $(I + K_k)$ with the help of the auxiliary function G_k as

$$I + K_k = Q \,\mathrm{diag}\, \Big(G_k(\lambda_1^{(k)}), \ldots, G_k(\lambda_r^{(k)}), 2, \ldots, 2\Big)\; Q^t \ . \tag{40}$$

Also from (25) it follows that

$$A^\dagger = Q \,\mathrm{diag}\, \left(\frac{1}{\lambda_1^{(0)}}, \ldots, \frac{1}{\lambda_r^{(0)}}, 0, \ldots, 0\right) Q^t \ . \tag{41}$$

Thus, we have, see e.g. [Bjö96],

$$P_{N(A)} = I - P_{R(A)} = I - A^\dagger A = Q \,\mathrm{diag}\, (0, \ldots, 0, 1, \ldots, 1)\, Q^t \ , \tag{42}$$

which together with (40) gives us

$$(I + K_k) P_{N(A)} = 2 P_{N(A)} \ . \tag{43}$$

Using the splitting (37) of the right-hand side b and (30) we get

$$b^1 = (I + K_0) b^0 = (I + K_0) b = (I + K_0) P_{R(A)}(b) + (I + K_0) P_{N(A)}(b) \ .$$

Combining this with (38) and (43) leads to

$$b^1 = (I + K_0) Ax + 2 P_{N(A)}(b) = A_1 x + 2 P_{N(A)} \ .$$

By a recursive argument we then obtain

$$b^{(k)} = A_k x + 2^k P_{N(A)}(b), \ \forall k \geqslant 0 \ . \tag{44}$$

This together with the null-space identity of Corollary 3 gives us

$$A_k b^{(k)} = A_k A_k x = A_k^2 x, \ \forall k \geqslant 0 \ . \tag{45}$$

From the convergence properties of the IFKOBS algorithm, see Theorem 3, (45) and the definition of the Moore-Penrose pseudoinverse $A^\dagger$, see e.g. [BO71], we obtain

$$\lim_{k\to\infty} A_k b^{(k)} = \lim_{k\to\infty} A_k A_k x = (A^\dagger A)(A^\dagger A) x =$$

$$(A^\dagger A A^\dagger)(Ax) = A^\dagger A x = A^\dagger P_{R(A)}(b) = x_{LS}$$

and the proof is complete.

□

Remark 5 *The result in* (36) *does not contradict Theorem* 4*. Indeed, if* (29) *is consistent, both relations* (31) *and* (36) *hold.*

Remark 6 *In the case that* (29) *is inconsistent, relation* (31) *no longer holds. In fact, if* $P_{N(A)}(b) \neq 0$ *we obtain from* (44)

$$\lim_{k\to\infty} \|b^{(k)}\| = \infty \ .$$

Acknowledgments

The work of C. Popa is partially supported by the PNCDI INFOSOC Grant 131/2004. M. Mohr is a member of the DFG Junior Research Group: Inverse Problems in Piezoelectricity and its Applications (grant Ka 1778/1).

References

[Bjö96] A. Björck. *Numerical Methods for Least Squares Problems*. SIAM, 1996.

[BO71] L. T. Boullion and P. L. Odell. *Generalized inverse matrices*. Wiley-Interscience, New York, 1971.

[EP01] D. J. Evans and C. Popa. Projections and preconditioning for inconsistent least-squares problems. *International Journal on Computer Mathematics*, 78(4):599–616, 2001.

[GvL96] G. H. Golub and C. F. van Loan. *Matrix Computations*. The John Hopkins University Press, 1996. 3rd edition.

[Han98] P. C. Hansen. *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*. SIAM Monographs on Mathematical Modeling and Computation. SIAM, 1998.

[Kov70] S. Kovarik. Some iterative methods for improving orthogonality. *SIAM Journal on Numerical Analysis*, 7(3):386–389, 1970.

[MPR03a] M. Mohr, C. Popa, and U. Rüde. An Iterative Algorithm for Approximate Orthogonalisation of Symmetric Matrices. Technical Report 03–2, Lehrstuhl für Informatik 10 (Systemsimulation), Friedrich-Alexander-Universität Erlangen-Nürnberg, 2003.

[MPR03b] M. Mohr, C. Popa, and U. Rüde. Numerical Solution of Least-Squares Problems by a Modified Kovarik Algorithm. Technical Report 03–7, Lehrstuhl für Informatik 10 (Systemsimulation), Friedrich-Alexander-Universität Erlangen-Nürnberg, 2003.

[MPR04] M. Mohr, C. Popa, and U. Rüde. An Iterative Algorithm for Approximate Orthogonalisation of Symmetric Matrices. *International Journal of Computer Mathematics*, 81(2):215–226, 2004. see also [MPR03a].