МАТЕМАТИКА И МАТЕМАТИЧЕСКО ОБРАЗОВАНИЕ, 2002
MATHEMATICS AND EDUCATION IN MATHEMATICS, 2002
Proceedings of Thirty First Spring Conference of the Union of Bulgarian Mathematicians
Borovets, April 3–6, 2002

SOME DEVELOPMENTS OF KOVARIK'S APPROXIMATE ORTHOGONALIZATION METHODS

Constantin Popa

In this paper we construct new versions of the two Kovarik's approximate orthogonalization algorithms, for the case of symmetric and positive semi-definite real matrices. We prove that these algorithms generate sequences of matrices convergent to a matrix of the same kind, but having a smaller generalized spectral condition number. Some numerical experiments confirming these results are also presented for the normal equation associated to a collocation discretization of an integral equation of the first kind.

1. New KOVARIK-like algorithms. Let A be an $m \times n$ real matrix. We shall denote by $\sigma(A), \rho(A), (A)_i$ and A^t the spectrum, spectral radius, i-th row and the transpose of A (this last one with respect to the Euclidean scalar product and the associated norm, denoted by $\langle \cdot, \cdot \rangle, \| \cdot \|$, respectively). All the vectors that will appear will be considered as column vectors. The notation $\| A \|$ will be used for the spectral norm of B and is defined by

$$\| A \| = \sqrt{\rho(A^t A)} = \sqrt{\rho(A A^t)}. \tag{1}$$

Let $(a_k)_{k \geq 0}$ be the sequence of real numbers defined by $a_k = \dfrac{1}{2^{2k}} \dfrac{(2k)!}{(k!)^2}, k \geq 0$ and $(q_k)_{k \geq 0}$ an apriori fixed sequence of positive integers. Then, the two approximate orthogonalization algorithms, firstly proposed by Z. Kovarik in [4] and extended by the author in [5] and [6], can be written as follows.

Algorithm (KOA) Set $A_0 = A$ and for $k \geq 0$ do

$$H_k = I - A_k A_k^t,\ \Gamma_k = I + a_1 H_k + \ldots + a_{q_k} H_k^{q_k},\ A_{k+1} = \Gamma_k A_k. \tag{2}$$

Algorithm (KOB) set $A_0 = A$ and for $k \geq 0$ do

$$K_k = (I - A_k A_k^t)(I + A_k A_k^t)^{-1},\ \Gamma_k = I + K_k,\ A_{k+1} = \Gamma_k A_k. \tag{3}$$

The following results were proved in [5] and [6].

Theorem 1. *Let us suppose that*

$$\| A A^t \| = \rho(A A^t) = \rho(A^t A) < 1. \tag{4}$$

Then, the sequence $(A_k)_{k \geq 0}$ generated by any of the previous algorithms **(KOA)** *or* **(KOB)** *converges to the matrix*

$$A_\infty = ((A A^t)^{\frac{1}{2}})^+ A, \tag{5}$$

where by B^+ we denoted the Moore-Penrose pseudoinverse of the matrix B (see e.g. [1]*).*

Let us now suppose that $m = n$ and the matrix A is symmetric and positive semi-definite, i.e.

$$A = A^t, \ \ \langle Ax, x\rangle \geq 0, \ \forall x \in \mathbb{R}^n. \tag{6}$$

Then, we consider the following simplified versions of the above methods.

Algorithm (KOAS) Set $A_0 = A$ and for $k \geq 0$ do

$$H_k = I - A_k, \ \Gamma_k = I + a_1 H_k + \ldots + a_{q_k} H_k^{q_k}, \ A_{k+1} = \Gamma_k A_k. \tag{7}$$

Algorithm (KOBS) Set $A_0 = A$ and for $k \geq 0$ do

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

The following convergence result holds.

Theorem 2. *Let us suppose that*

$$\parallel A \parallel = \rho(A) = \rho(A^t) < 1. \tag{9}$$

Then, the sequence $(A_k)_{k\geq 0}$ generated by any of the previous algorithms **(KOAS)** *or* **(KOBS)** *converges to the matrix*

$$\hat{A}_\infty = (A^{\frac{1}{2}})^+ A. \tag{10}$$

Proof. From (6) it follows that it exists an orthonormal $n \times n$ matrix Q such that (see e.g. [1])

$$Q^t A Q = D_0 = diag(\sigma_1, \sigma_2, \ldots, \sigma_r, 0, \ldots, 0), \tag{11}$$

$$\sigma(A) = \{\sigma_1, \sigma_2, \ldots, \sigma_r, 0\}, \ \sigma_i > 0, \ i = 1, \ldots, r, \tag{12}$$

where $r \leq n$ is the rank of the matrix A. Moreover, from (9) we obtain

$$0 < \sigma_i < 1, \ \forall i = 1, \ldots, r. \tag{13}$$

Then, for the algorithm **(KOAS)** the proof of convergence follows exactly the same steps as in [6].
For the algorithm **(KOBS)**, using (8), (11) and the orthonormality of the matrix Q we obtain

$$A_{k+1} = Q D_{k+1} Q^t \tag{14}$$

and

$$D_{k+1} = [I + (I - D_k)(I + D_k)^{-1}] D_k, \forall k \geq 0. \tag{15}$$

Because D_0 is a diagonal matrix, so will be $D_k, \forall k \geq 0$. Moreover, $D_k = diag(\sigma_1^{(k)}, \ldots, \sigma_r^{(k)}, 0, \ldots, 0)$ with $\sigma_j^{(k)} > 0, \forall j = 1, \ldots, r$ recursively defined by

$$\sigma_j^{(k+1)} = \frac{2\sigma_j^{(k)}}{1 + \sigma_j^{(k)}}, \ k \geq 0, \ \sigma^{(0)} = \sigma_j \in (0, 1). \tag{16}$$

By analyzing the function $x \rightarrow \dfrac{2x}{1+x}, x \in (0, 1)$, we obtain that all the sequences $(\sigma_j^{(k)})_{k\geq 0}, j = 1, \ldots, r$ converge to 1. Then, from (14) we get

$$\lim_{k\to\infty} A_k = Q \ diag(1, 1, \ldots, 1, 0, \ldots, 0) \ Q^t. \tag{17}$$

Thus, as in [5] we obtain the equality $\hat{A}_\infty = Q\ diag(1,1,\ldots,1,0,\ldots,0)\ Q^t$, which completes the proof. □

Remark 1. Conditions of the type (4) or (9) can be obtained by an appropriate scaling of the entries of A (see e.g. the next section of this paper).

Remark 2. We have to observe that the convergence results from the above Theorem 2 do not hold directly from those in Theorem 1. Indeed, if A is as in (6), then so are all the matrices A_k thus, their square roots $A_k^{\frac{1}{2}}$ exist with the properties

$$A_k^{\frac{1}{2}} = (A_k^{\frac{1}{2}})^t, A_k = A_k^{\frac{1}{2}} A_k^{\frac{1}{2}} = A_k^{\frac{1}{2}} (A_k^{\frac{1}{2}})^t. \tag{18}$$

But, in the construction of A_{k+1} in (7) or (8) the whole matrix A_k is used and not only $A_k^{\frac{1}{2}}$.

Remark 3. If $\tilde{D} = diag(1,1,\ldots,1,0,\ldots,0)$ is from (17) the following result holds with respect to the "quasi-orthonormality" of the rows of $\hat{A}_\infty$

$$\langle (\hat{A}_\infty)_i, (\hat{A}_\infty)_j \rangle = \langle \tilde{D}(Q)_i, (Q)_j \rangle,\ \forall i,j = 1,\ldots,n \tag{19}$$

(for the proof see [5]).

The equalities (19) tell us that, if A would be invertible, then $\tilde{D} = I$ and $\hat{A}_\infty$ would have orthonormal rows. In the general case, we can obtain only an improvement of the values of the angles between the rows of A (for $\hat{A}_\infty$ these values become "closer" to 90°; see the numerical experiments in [5] and [6]). We also obtain an improvement of the generalized spectral condition number of A, defined by (see also (11))

$$k_2(A) = \frac{\max\{\sigma_j, 1 \le j \le r\}}{\min\{\sigma_j, 1 \le j \le r\}}, \tag{20}$$

which for the matrix $\hat{A}_\infty$ from (10) obviously has the "ideal" value

$$k_2(\hat{A}_\infty) = 1. \tag{21}$$

2. Applications to normal equations. We come back now to the general case of a rectangular $m \times n$ real matrix $B, m \neq n$ and, for $c \in I\!R^m$ we consider the linear least-squares problem: find $x^* \in I\!R^n$ such that

$$\| Bx^* - c \| = \min\{\| Bx - c \|, x \in I\!R^n\}. \tag{22}$$

It is well known (see e.g. [1]) that (22) is equivalent with the associated normal equation problem: find $x^* \in I\!R^n$ such that

$$B^t B x^* = B^t c. \tag{23}$$

If we denote by $\tilde{A}$ the $n \times n$ matrix from (23), i.e.

$$\tilde{A} = B^t B \tag{24}$$

and we define

$$A = \frac{1}{\| \tilde{A} \|_\infty} \tilde{A} \tag{25}$$

(with $\| \tilde{A} \|_\infty = \max_{1 \le i \le n} \sum_{j=1}^n |(\tilde{A}_{ij}|)$ we obtain that A is a symmetric and positive semi-definite matrix, also satisfying the assumption (9). Thus, by applying few steps of one of the algorithms **(KOAS)** or **(KOBS)** to the matrix A we shall obtain an approximation of $\hat{A}_\infty$ having a better condition number (defined by (20)) than the original matrix A.

This can much improve the behavior of both direct and iterative methods for solving (22) through (23).

Remark 4. In order to don't modify the set of solutions for (22) (or (23)) we must apply transformations similar to (7) and (8) also to the right hand side of (23) (see [3]), i.e. $b^0 = B^t c$ and

(26) $$H_k = I - A_k,\ \Gamma_k = I + a_1 H_k + \ldots + a_{q_k} H_k^{q_k},\ b^{k+1} = \Gamma_k b^k.$$

(27) $$K_k = (I - A_k)(I + A_k)^{-1}, \Gamma_k = I + K_k,\ b^{k+1} = \Gamma_k b^k.$$

We considered in our tests the first kind integral equation: find $x \in L^2([0,1])$ such that

(28) $$Tx(s) = \int_0^1 k(s,t)x(t)dt = y(s),\ s \in [0,1],$$

with the elements

(29) $$y(s) = s \quad , \quad k(s,t) = \frac{1}{\sqrt{1+s^2t^2}}, \quad s,t \in [0,1].$$

Then, for given m and n and the points $s_i, i = 1, \ldots, m$ and $\tau_j, j = 1, \ldots, n$, defined by

(30) $$s_i = \frac{i-1}{m-1} \quad , \quad \tau_j = \frac{j-1}{n-1}.$$

we discretized the problem (27) following the collocation technique described in [2]. In this way we obtained a least-squares problem as (22), with a full rank but ill-conditioned matrix B. We constructed the matrix A as in (24)-(25) and applied to it the algorithm **(KOBS)**. In Table 1 we present (for the case m=16, n=8) the values of $k_2(A_k)$ for different values of k. Then we fixed $k_2(A_k) = 1.15$ (see the eight column in Table 1) and, for different values for m and n we determined the numbers of iteration for obtaining it. The corresponding values are presented in Table 2.

Remark 5. We observe the good improvement of $k_2(A_k)$ in Table 1 and also the fact that the numbers of iterations from Table 2 are (almost) independent on the dimensions m and n. Similar results were obtained with the algorithm **(KOAS)**, with $q_k = N$, $\forall k \geq 0$ and different values of $N \geq 1$. All the numerical experiments have been performed with the numerical linear algebra software **OCTAVE**, free available on the Internet.

Table 1.

k	0	10	20	30	40	50	**60**	70
$k_2(A_k)$	10^{16}	10^{14}	10^{11}	10^8	10^5	10^3	**1.15**	1.0001

Table 2.

m	n	Number of iterations
16	8	61
32	16	60
64	32	62
128	64	62
256	128	63

REFERENCES

[1] A. Björck. Numerical methods for least squares problems, SIAM Philadelphia, 1996.
[2] H. W. Engl. Regularization methods for the stable solution of inverse problems. *Surv. Math. Ind.* **3** (1993), 71–143.
[3] D. J. Evans, C. Popa. Projections and preconditioning for inconsistent least-squares problems, *Intern. J. Comp. Math.* **78** (2001), 599–616.
[4] Z. Kovarik. Some iterative methods for improving orthonormality. *SIAM J. Numer. Anal.* **7** (1970), 386–389.
[5] C. Popa. A method for improving orthogonality of rows and columns of matrices. *Intern. J. Comp. Math.* **77** (2001), 469–480.
[6] C. Popa. Extension of an approximate orthogonalization algorithm to arbitrary rectangular matrices. *Linear Alg. Appl.* **331** (2001), 181–192.

Constantin Popa
Faculty of Mathematics and Computer Science
“OVIDIUS” University of Constanta
Blvd. Mamaia 124, 8700 Constanta, Romania
e-mail: cpopa@univ-ovidius.ro

НОВИ ВЕРСИИ НА МЕТОДИ НА КОВАРИК ЗА ПРИБЛИЖЕНА ОРТОГОНАЛИЗАЦИЯ

Константин Попа

Конструирани са нови версии на два алгоритъма на Коварик за приближена ортогонализация на симетрични и положително полуопределени реални матрици. Доказано е, че тези алгоритми пораждат редици от матрици, сходящи към матрица от същия тип, но с по-малко обобщено спектрално число на обусловеност. Представени са числени експерименти, които потвърждават тези резултати в случая на нормално уравнение, получено от колокационна дискретизация за интегрално уравнение от първи ред.