TWMS J. Pure Appl. Math., V.4, N.2, 2013, pp.115-124

ON A PARALLEL COMPUTATION METHOD FOR SOLVING LINEAR ALGEBRAIC SYSTEM WITH ILL-CONDITIONED MATRIX

M. OTELBAEV[1], D. ZHUSUPOVA[1], B. TULEUOV[1]

ABSTRACT. A new method of finding approximate solutions of linear algebraic systems with ill-conditioned or singular matrices is presented. This method can effectively be used for arranging parallel computations for matrices of large size. Difference of the method suggested from the known is that existence of zero eigenvalues of the matrix of system doesn't influence by no means efficiency of iterative process. Only small but nonzero singular values of the matrix are important.

Keywords: ill conditioned matrices, eigenvalues, approximate solutions, parallel computation.

AMS Subject Classification: 68W30.

1. Introduction

This work is to continue [1], where we have considered the equation

$$Ax = f. \tag{1}$$

Here A is a quadratic matrix of order n and f is n-dimensional vector. In [1] the problem of parallel computation for solving equation (1) has been considered and effective parallel algorithm has been developed for the matrix A with bounded inverse.

In this paper we suggest a method for finding and parallel computation algorithm for the approximate solutions of the problem (1), when matrix A is noninvertible or ill-conditioned. Raising effeciency of solving of the large system of linear equations dependes on development of the high-effective calculating techniques. Now multiprocessor systems and supercomputers is highly developed. Distribution of calculations into parallel branches implies the increase of solving of the general problem. Parallel computation of linear algebraic problems have been considered, for example, in monographs [2], [5], [8], and software realization questions in [3].

The difference of the offered method from the known ones consists of the existence of zero eigenvalues of the matrix A, doesn't influence efficiency of iterative process in any way. Only small, but nonzero eigenvalues of A^*A are important. Besides, estimates obtained here in the Theorem 3 for the solution does not depend on small and nonzero eigenvalues of the matrix $A^* \times A$. Offered parallelizing process for the linear algebraic system with an arbitrary matrix when using of k computers (processes) reduces time expenses approximately k times. The main known methods are applied to cases of the band and sparse matrices. But even in these cases the effect received by us wasn't reached.

2. Problem statement and main results

We denote by A^* adjoint matrix of A. Nonnegative square roots of the eigenvalues of the nonnegative matrix A^*A we denote by $s_j(A) = s_j (j = 1, 2, ...)$ and numerate them in non-increasing order taking into account multiplisities. Orthonormal eigenvectors of the operator A^*A corresponding to s_j^2 we write as e_j ($j = 1, 2, ...; A^*Ae_j = s_j^2 e_j$).

Numbers $s_1(A) \geq s_2(A) \geq ... \geq s_n(A)$ are called singular numbers of the matrix A.

[1]L.N. Gumilev Eurasian National University, Astana, Kazakhstan,
e-mail: otelbaevm@mail.ru, zhus_dinara@mail.ru, berik_t@yahoo.com
Manuscript received July 2011.

Note that the notion "ill-conditioned matrix" is relative, which often depends on hardware capabilities and may be compensed by the increase of the possibilities of the computing techniques. We, roughly speaking, say that the matrix A is ill-conditioned if some of its singular numbers are small enough or zero (indeed, it is possible to show easily that this definition is equivalent to the traditional one, if take as norm $\|A\| = \sup_i s_i$).

Further vector's Euclid norm and modulus of the number we write as $|\cdot|$, and operator's norm of matrix as $\|\cdot\|$, and scalar composition as $\langle\cdot,\cdot\rangle$.

Let A and f be from (1) and for $\varepsilon \geq 0$ consider the functional

$$J_\varepsilon(x) = |Ax - f|^2 + \varepsilon|x|^2.$$

We will find $\mathring{x}$, which is a solution of the problem

$$\inf J_\varepsilon(x) = J_\varepsilon(\mathring{x}). \tag{2}$$

In the left-hand side of (2) *infimum* is taken with respect to all vectors $x \in R^n$. Since unit ball in R^n is compact, solution of (2) exists.

Remark 2.1. *If matrix A is invertible, then for $\varepsilon = 0$ problem (2) has the unique solution $\mathring{x} = A^{-1}f$. If A is noninvertible, then $A\mathring{x}$ gives the best approximation of f by the elements Ax. If $\varepsilon \neq 0$ and A is invertible, then $\mathring{x}$ is the approximate solution of the equation $Ax = f$.*

If $\varepsilon = 0$ and matrix A is noninvertible, then the problem (1) may have several solutions. In this case we search for solution with minimal norm.

Lemma 2.1. *If $\varepsilon \geq 0$ and $\mathring{x}$ is a solution of (2), then $A^*(A\mathring{x} - f) + \varepsilon\mathring{x} = 0$.*

Proof. Let $\mathring{x}$ be a solution of (2) and $\omega = A^*(A\mathring{x} - f) + \varepsilon\mathring{x} \neq 0$. Considering $J_\varepsilon(\mathring{x} + \delta\omega)$ we have:

$$J_\varepsilon(\mathring{x} + \delta\omega) = J_\varepsilon(\mathring{x}) + 2\delta\langle A^*(A\mathring{x} - f) + \varepsilon\mathring{x}, \omega\rangle + \delta^2(|A\omega|^2 + \varepsilon|\omega|^2) =$$

$$= J_\varepsilon(\mathring{x}) + 2\delta|\omega|^2 + \delta^2(|A\omega|^2 + \varepsilon|\omega|^2).$$

Let a number δ to satisfy the following conditions

$$\delta < 0, -2\delta < \delta^2 \frac{|A\omega|^2 + \varepsilon|\omega|^2}{|\omega|^2}.$$

Such a choice is possible by assumption $\omega \equiv A^*(A\mathring{x} - f) + \varepsilon\mathring{x} \neq 0$. Then we get $J_\varepsilon(\mathring{x} + \delta\omega) < J_\varepsilon(\mathring{x})$ that is a contradiction. □

Lemma 2.2. *Let $\varepsilon \geq 0$ and $\mathring{x}$ be a solution of (2). Then for all $x \in H$ we have*

$$\varepsilon x + A^*(Ax - f) = (\varepsilon + A^*A)(x - \mathring{x}).$$

Proof. Using Lemma 2.1 we obtain

$$\varepsilon x + A^*(Ax - f) = \varepsilon x + A^*(Ax - f) - \varepsilon\mathring{x} - A^*(A\mathring{x} - f) =$$

$$= \varepsilon(x - \mathring{x}) + A^*A(x - \mathring{x}) = (\varepsilon + A^*A)(x - \mathring{x}).$$

□

Now we define the sequence x_j $(j = 1, 2, ...)$ by the following formula

$$x_j = \delta \sum_{k=0}^{j-1} [E - \delta(A^*A + \varepsilon E)]^k A^* f, \tag{3}$$

where δ satisfies the condition

$$0 < \delta < \frac{2}{\|A^*A\| + \varepsilon}. \tag{4}$$

Theorem 2.1. *Let $\varepsilon \geq 0$, δ be given by (4) and $\mathring{x}$ be a solution of (2), x_j be constructed by (3). Then*

$$x_j - \mathring{x} = -[E - \delta(A^*A + \varepsilon E)]^j \mathring{x}, \tag{5}$$

and x_j converges to $\mathring{x}$ as $j \to +\infty$ at the geometric rate, i.e. there exists $\rho > 0$ and

$$|x_j - \mathring{x}| \leq C \cdot \rho^j, \tag{6}$$

where C is a constant which depends on δ and ε.

Proof. By using Lemma 2.1 we have

$$A^*f = \varepsilon\mathring{x} + A^*A\mathring{x}.$$

Substituting A^*f into (3) we get

$$x_j = \delta \sum_{k=0}^{j-1} [E - \delta(A^*A + \varepsilon E)]^k [\varepsilon\mathring{x} + A^*A\mathring{x}] =$$

$$= \sum_{k=0}^{j-1} [E - \delta(A^*A + \varepsilon E)]^k [E - E + \delta(\varepsilon + A^*A)]\mathring{x} =$$

$$= -\sum_{k=1}^{j} [E - \delta(A^*A + \varepsilon)]^k \mathring{x} + \sum_{k=0}^{j-1} [E - \delta(A^*A + \varepsilon E)]^k \mathring{x} = -(E - \delta(A^*A + \varepsilon))^j \mathring{x} + \mathring{x}.$$

This implies (5).

Further, since the matrix $E - \delta(A^*A + \varepsilon E)$ is self-adjoint, its norm is equal to the maximum of modulus of eigenvalues. These eigenvalues indeed are $1 - \delta({s_j}^2 + \varepsilon)$, $(j = 1, 2, ..., n)$. If for each $j = 1, 2, ..., n$ eigenvalues satisfy

$$-1 < 1 - \delta({s_j}^2 + \varepsilon) < 1,$$

then we get

$$\|E - \delta(A^*A + \varepsilon E)\| < 1. \tag{7}$$

These inequalities hold, if conditions $\delta(\max\limits_{j=1,2,...,n} {s_j}^2 + \varepsilon) < 2$ and $\delta > 0$ take place. But $\max\limits_{j=1,2,...,n} {s_j}^2 = \|A^*A\|$. From the condition (4) follows (7) and by (7) we get (6). □

Note that results similar to Theorem 2.1 for linear ill-posed problems have been obtained in [4] (see [4], p. 238).

We denote the space generated by eigenvectors of the matrix A^*A corresponding to zero eigenvalues by $R_0^{(n)}$, i. e. if $\mathring{x} \in R_0^{(n)}$ then $x = \sum_{k=j_0}^{n} x_j e_j$ and $A^*Ae_k = 0$ for $k = j_0, ..., n$. $R_0^{(n)}$ is the kernel of matrix A^*A.

If matrix A is invertible, then the space $R_0^{(n)}$ is empty.

Lemma 2.3. *If $x \in R_0^{(n)}$ then $\langle A^*f, x\rangle = 0$, i.e. A^*f belongs to $R^{(n)} \ominus R_0^{(n)}$ which is the orthogonal complement of $R_0^{(n)}$.*

Proof. For $\varepsilon = 0$ by using the Lemmas 2.1 and 2.2 we obtain $A^*f = A^*A\mathring{x}$. Let $x \in R_0^{(n)}$, then

$$\langle A^*f, x\rangle = \langle A^*A\mathring{x}, x\rangle = \langle \mathring{x}, A^*Ax\rangle = 0.$$

□

Lemma 2.4. *If $\varepsilon > 0$ and $\mathring{x}$ is a solution of (2), $x \in R_0^{(n)}$, then $\langle \mathring{x}, x\rangle = 0$, i.e. $\mathring{x}$ belongs to $R^{(n)} \ominus R_0^{(n)}$.*

Proof. For $\varepsilon > 0$ using Lemmas 2.1 and 2.2 we have

$$\varepsilon\langle \mathring{x}, x\rangle = \langle A^*f, x\rangle = \langle A^*A\mathring{x}, x\rangle = -\langle \mathring{x}, A^*Ax\rangle = 0.$$

For $\varepsilon > 0$ this implies the lemma. □

Note that for $\varepsilon = 0$ the solution of (2) is determined up to a term, which is a solution of the equation $Ax = 0$, but sequence x_j $(j = 1, 2, ...)$ by Lemma 2.2 converges to the solution of (2) belonging to $R^{(n)} \ominus R_0^{(n)}$. In further for $\varepsilon = 0$ we take as $\mathring{x}(0)$ the limit of the sequence x_j from (3).

Obviously the solution of (2) depends on ε. So sometimes we write $\mathring{x} = \mathring{x}(\varepsilon)$.

We have

Lemma 2.5. *If $\mathring{x}(0)$ is a solution of (2), then for every $\varepsilon > 0$ and $\delta \geq 0$*

$$\mathring{x}(\varepsilon) = (A^*A + \varepsilon E)^{-1}A^*A\mathring{x}(o) = (A^*A + \varepsilon E)^{-1}A^*f,$$
$$\mathring{x}(0) = (E + \varepsilon(A^*A)^{-1})\mathring{x}(\varepsilon),$$
$$\mathring{x}(\varepsilon) = (A^*A + \varepsilon E)^{-1}(A^*A + \delta E)\mathring{x}(\delta),$$
$$\mathring{x}(\varepsilon) - \mathring{x}(\delta) = (\delta - \varepsilon)(A^*A + \varepsilon E)^{-1}\mathring{x}(\delta).$$

Proof. By Lemma 2.1 for each ε, $\delta \geq 0$ we obtain

$$(A^*A + \varepsilon E)\mathring{x}(\varepsilon) = (A^*A + \delta E)\mathring{x}(\delta).$$

This implies assertion of Lemma 2.5. □

From proved lemmas and Theorem 2.1 we state

Theorem 2.2. *a) The solution $\mathring{x}(\varepsilon)$ of (2) continuously depends on $\varepsilon > 0$ and $\mathring{x}(\varepsilon) = (A^*A + \varepsilon E)^{-1}A^*f$.*

b) For $j \to +\infty$ the limit of sequence $x_j(\varepsilon)$ from (3) continuously depends on $\varepsilon \geq 0$.

c) If $s_1 \geq s_2 \geq \ldots \geq s_{j_0} > 0, s_{j_0+1} = s_{j_0+2} = ... = 0$ are the eigenvalues of the matrix A^*A and e_1, e_2, ..., e_n are corresponding orthonormal system of eigenvectors, $\mathring{x}(\varepsilon)$ $(\varepsilon > 0)$ is a solution of (2) and $x_j(\varepsilon)$ is from (3), then $\mathring{x}(\varepsilon)$, $x_j(\varepsilon) \in R^{(n)} \ominus R_0^{(n)}$, $(j = 1, 2, \ldots)$

$$\begin{array}{c} x_{jk}(\varepsilon) - \mathring{x}_k(\varepsilon) = (1 - \delta(s_k^2 + \varepsilon))^j \mathring{x}_k(0) \qquad by \qquad 1 \leq k \leq j_0, \\ x_{jk}(\varepsilon) = \mathring{x}_k(\varepsilon) \qquad by \qquad j_0 + 1 \leq k. \end{array} \tag{8}$$

Here $x_{jk}(\varepsilon) = \langle x_j(\varepsilon), e_k\rangle$, $\mathring{x}_k(\varepsilon) = \langle \mathring{x}(\varepsilon), e_k\rangle$.

d) Number $\rho > 0$ from Theorem 2.1 is defined by

$$\rho = \min\{(1 - \delta(s_{j_0}^2 + \varepsilon)), (1 - \delta(\|A^*A\| + \varepsilon))\} < 1.$$

Note that if the matrix A^*A hasn't zero eigenvalues, then j_0 is taken as n.

The item c) of Theorem 2.2 implies that the vector $\mathring{x}(\varepsilon)$ for $\varepsilon = 0$ has minimal norm among all solutions of problem 1. Furthermore, for each $\varepsilon \geq 0$ $\mathring{x}(\varepsilon)$ and $x_j(\varepsilon)$ $(j = 1, 2, \ldots)$ belong to subspace $R^{(n)} \ominus R_0^{(n)}$, where $R_0^{(n)}$ is the kernel of matrix A^*A.

3. Parallelization

Below we suggest a method of parallel computation for solving the problem (2) based on Theorems 2.1 and 2.2.

Let n be large enough integer and we have $N+1$-processor system. Let k_0, k_1, $\ldots$, k_N be such integers that $k_{m-1} + 1 < k_m$, $m = 0, 1, \ldots, N$, $k_0 = 0$, $k_N = n$. We define matrices A_m and $(A^*)_m$, $m = 1, 2, \ldots, N$ as follows

$$A_m = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 \\ a_{k_{m-1}+1,1} & a_{k_{m-1}+1,2} & a_{k_{m-1}+1,3} & \ldots & a_{k_{m-1}+1,n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{k_m,1} & a_{k_m,2} & a_{k_m,3} & \ldots & a_{k_m,n} \\ 0 & 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix},$$

$$(A^*)_m = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 \\ \tilde{a}_{k_{m-1}+1,1} & \tilde{a}_{k_{m-1}+1,2} & \tilde{a}_{k_{m-1}+1,3} & \dots & \tilde{a}_{k_{m-1}+1,n} \\ \dots & \dots & \dots & \dots & \dots \\ \tilde{a}_{k_m,1} & \tilde{a}_{k_m,2} & \tilde{a}_{k_m,3} & \dots & \tilde{a}_{k_m,n} \\ 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix} (m = 1, 2, \dots, N).$$

The lines numerated from $k_{m-1}+1$ to k_m of these matrices coincide. Here $\tilde{a_{kj}}$ and a_{kj} are elements of A^* and A such that $\tilde{a_{kj}} = a_{jk}$. Also we use vectors $\omega^j = [E - \delta(A^*A + \varepsilon E)]\,\omega^{j-1}$, $\omega^0 = \delta A^* f$, $j = 1, 2, \dots$. Then formula (3) can be written in the following way

$$x_{j+1} = x_j + \omega^j, x_1 = \omega^0, j = 1, 2, \dots$$

The matrices A_m and $(A^*)_m$ are passed to processors C_m $(m = 1, 2, \dots, N)$ before computation and vector $\omega^0 = \delta A^* f$ to root processor C_{N+1}.

(1) Processor C_{N+1} forms j-th approximation of x_j and passes vector ω^{j-1} to processors C_m $(m = 1, 2, \dots, N)$. Each processor C_m calculates $A_m\omega^{j-1}$ spending $(k_m - k_{m-1})n$ multiplications, $(k_m - k_{m-1})(n-1)$ additions and sends vector to C_{N+1}.

(2) C_{N+1} forms vector $A\omega^{j-1} = \sum\limits_{m=1}^{N} A_m\omega^{j-1}$ and spends $(n-1)N$ additions. C_{N+1} transmits vector $A\omega^{j-1}$.

(3) Processor C_m calculates $(A^*)_m A\omega^{j-1}$ and sends it to C_{N+1}. C_m spends $(k_m - k_{m-1})n$ multiplications and $(k_m - k_{m-1})(n-1)$ additions. C_m transmits $(A^*)_m A\omega^{j-1}$ to processor C_{N+1}.

(4) Summing up recieved vectors C_{N+1} gets $A^*(A\omega^{j-1}) = \sum\limits_{m=1}^{N} (A^*)_m A\omega^{j-1}$. Root processor calculates $\omega^j = (1 - \delta\varepsilon)\omega^{j-1} - \delta(A^*)_m A\omega^{j-1}$ and forms approximate solution $x_{j+1} = x_j + \omega^j$, $j = 1, 2, \dots$. It spends $2n$ multiplications and $(n-1)N + 2n$ additions.

Amount of operations per cycle which root processor C_{N+1} carries out consists of $2n$ multiplications and $2(n-1)N + 2n$ additions. Each processor C_m $(m = 1, 2, \dots, N)$ spends $2(k_m - k_{m-1})n$ multiplications and $2(k_m - k_{m-1})(n-1)$ additions per cycle. All processors spend $2n^2$ multiplications and $2n(n-1)$ additions per cycle. Since s iterations all computers spend $2sn^2 + n^2 + n$ multiplications and $2sn(n-1) + n(n-1)$ additions. Division is absent.

4. Some discussions

It follows from Theorem 2.2 that for efficiency of the iteration formula (3) the existence of small but nonzero eigenvalues of matrix A^*A is important, but zero eigenvalues aren't so important! Therefore we come to the question: Can it reduce "noises" due to nonzero small eigenvalues of matrix A^*A? It turns out that it is possible. We describe it by easy example.

Let $\varepsilon = 0.01$ and

$$A = \begin{pmatrix} 1 & 1 \\ 3 & 3.001 \end{pmatrix}, \ f = \begin{pmatrix} 2 \\ 6.006 \end{pmatrix}. \tag{9}$$

Matrix

$$A^*A = \begin{pmatrix} 1 & 3 \\ 1 & 3.001 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 3 & 3.001 \end{pmatrix}$$

has a small nonzero eigenvalue. Iterative process by formula (8) may last long. However, if matrix A is replaced with its approximation

$$\tilde{A} = \begin{pmatrix} 1 & 1 \\ 3 & 3 \end{pmatrix}, \ \tilde{f} = \begin{pmatrix} 2 \\ 6 \end{pmatrix}, \tag{10}$$

then we have

$$\tilde{A}^*\tilde{A} = \begin{pmatrix} 1 & 3 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 3 & 3 \end{pmatrix} = \begin{pmatrix} 10 & 10 \\ 10 & 10 \end{pmatrix}.$$

The eigenvalues of this matrix are equal to $\lambda_1^2 = 20$, $\lambda_2^2 = 0$ and its norm is 20. So δ may be taken from the interval $(0, \frac{1}{10})$. Let's take $\delta = \frac{1}{20} < \frac{1}{10}$. Then by Theorem 2.2 we obtain $\rho = 0$. Therefore the problem 1 with matrix $\tilde{A}$ and vector $\tilde{f}$ from (10) is solved in one step.

The solution of problem 1 is the vector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Equation

$$\tilde{A}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \end{pmatrix}$$

has a solution vector $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ such that $x_1 + x_2 = 2$. Vector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ satisfies this condition and has a minimal norm among all vectors. The found vector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ will be the approximate solution of (2) with matrix A and vector f from (9). Indeed

$$\begin{pmatrix} 1 & 1 \\ 3 & 3.001 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 2 \\ 6.006 \end{pmatrix} = \begin{pmatrix} 0 \\ 0.006 \end{pmatrix}.$$

We have $|A\mathring{x} - f| = 0.006 \approx 0$. (Recall that we try to reduce the norm $|Ax - f|$ increasing the norm $|x|$ not too much).

The vector $\tilde{x} = \begin{pmatrix} \tilde{x_1} \\ \tilde{x_2} \end{pmatrix} = \begin{pmatrix} -4 \\ 6 \end{pmatrix}$ is the actual solution of the system

$$\begin{pmatrix} 1 & 1 \\ 3 & 3.001 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 6.006 \end{pmatrix}.$$

For $\mathring{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ we get

$$|A\mathring{x} - f|^2 + \varepsilon|\mathring{x}|^2 = \left|\begin{pmatrix} 0 \\ 0.006 \end{pmatrix}\right|^2 + 0.01\left|\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right|^2 = (0.006)^2 + 0.01 \approx 0.01.$$

And for $\begin{pmatrix} -4 \\ 6 \end{pmatrix}$

$$|A\tilde{x} - f|^2 + \varepsilon|\tilde{x}|^2 = 0 + 0.01(16 + 24) = 0.4.$$

Therefore, for $\varepsilon = 0.01$ the vector $\mathring{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is closer to the solution of problem 1 than $\tilde{x} = \begin{pmatrix} -4 \\ 6 \end{pmatrix}$.

This simple idea checked simple example that we will develop in the next work with matrices arisen in solving numerically ill-posed direct and inverse problems of mathematical physics.

In general, this effect is not always possible. But we have

Theorem 4.1. *Let $\varepsilon \geqslant 0$ and x_j $(j = 1, 2, \dots)$ be the sequence of vectors from (3). Then*

a) If $\gamma > 0$ and j satisfies $(1 - \delta(\gamma + \varepsilon))^{2j} \leqslant \gamma$, then we obtain the following inequality

$$|Ax_j(\varepsilon) - f| \leqslant 2|f|\sqrt{\gamma} + \gamma|\mathring{x}(\varepsilon)| + |A\mathring{x}(\varepsilon) - f|\,;$$

b) If j is given by

$$(1 - \delta(\gamma + \varepsilon))^{2j} \leqslant \gamma^2,$$

then

$$|A^*A\left(x_j(\varepsilon) - \mathring{x}(\varepsilon)\right)|^2 \leqslant \gamma^2\left[|A^*A\mathring{x}|^2 + |\mathring{x}|^2\right];$$

c) If j is given by

$$(1 - \delta(\gamma + \varepsilon))^{2j} \leqslant \frac{2}{5\,\|A^*\|}\gamma,$$

then

$$|A^*A\left(x_j(\varepsilon)-\mathring{x}(\varepsilon)\right)|^2 \leqslant 8\gamma\,|f|^2\,;$$

d) For $\varepsilon=0$, if j is given by

$$(1-\delta\gamma)^{2j} \leqslant \frac{2}{5\,\|A^*\|}\gamma,$$

then

$$|A^*A\left(x_j(0)-f\right)|^2 \leqslant 8\gamma\,|f|^2\,. \tag{11}$$

Proof. Let $\varepsilon \geqslant 0$ and

$$\inf_{\{x\}}\left(|Ax-f|^2+\varepsilon\,|x|^2\right)=|A\mathring{x}(\varepsilon)-f|^2+\varepsilon\,|\mathring{x}(\varepsilon)|^2\,.$$

For any vector u we have

$$|Au|^2=\langle Au,Au\rangle=\langle A^*Au,u\rangle=\left|(A^*A)^{\frac{1}{2}}\,u\right|^2.$$

Therefore, by assumption $(A^*A)^{\frac{1}{2}}\,e_k=s_ke_k$ and using (8) from Theorem 2.2 we have

$$\begin{aligned}|A\left(x_j(\varepsilon)-\mathring{x}(\varepsilon)\right)|^2 &= \left|(A^*A)^{\frac{1}{2}}\left(x_j(\varepsilon)-\mathring{x}(\varepsilon)\right)\right|^2=\sum_{k=1}^n s_k^2\left(x_{jk}(\varepsilon)-\mathring{x}_k(\varepsilon)\right)^2= \\ &= \sum_{k=1}^n s_k^2\left(1-\delta(s_k^2+\varepsilon)\right)^{2j}|\mathring{x}_k(\varepsilon)|^2\,.\end{aligned}$$

Hence, for all $\gamma>0$ we get

$$\begin{gathered}|A\left(x_j(\varepsilon)-\mathring{x}(\varepsilon)\right)|^2=\sum_{s_k^2>\gamma} s_k^2\left(1-\delta(s_k^2+\varepsilon)\right)^{2j}|\mathring{x}_k(\varepsilon)|^2+ \\ +\sum_{s_k^2\leqslant\gamma} s_k^2\left(1-\delta(s_k^2+\varepsilon)\right)^{2j}|\mathring{x}_k(\varepsilon)|^2\leqslant(1-\delta(\gamma+\varepsilon))^{2j}\sum_{s_k^2>\gamma}s_k^2\,|\mathring{x}_k(\varepsilon)|^2+ \\ +\gamma\sum_{s_k^2\leqslant\gamma}|\mathring{x}_k(\varepsilon)|^2\leqslant(1-\delta(\gamma+\varepsilon))^{2j}\sum_{k=1}^n s_k^2\,|\mathring{x}_k(\varepsilon)|^2+\gamma\sum_{k=1}^n|\mathring{x}_k(\varepsilon)|^2= \\ =(1-\delta(\gamma+\varepsilon))^j\left|(A^*A)^{\frac{1}{2}}\,\mathring{x}(\varepsilon)\right|^2+\gamma\,|\mathring{x}(\varepsilon)|^2= \\ =(1-\delta(\gamma+\varepsilon))^j\,|A\mathring{x}(\varepsilon)|^2+\gamma\,|\mathring{x}(\varepsilon)|^2\,.\end{gathered} \tag{12}$$

But

$$\begin{gathered}|A\mathring{x}(\varepsilon)|^2=|A\mathring{x}(\varepsilon)-f+f|^2\leqslant 2\left(|A\mathring{x}(\varepsilon)-f|^2+|f|^2\right)= \\ =2\left[\left(\inf_{\{x\}}\left(|Ax-f|^2+\varepsilon\,|x|^2\right)\right)+|f|^2\right]\leqslant \\ \leqslant 2(|f|^2+|f|^2)=4|f|^2.\end{gathered} \tag{13}$$

Using this estimate and (12) we arrive at the estimate

$$|A\left(x_j(\varepsilon)-\mathring{x}(\varepsilon)\right)|^2\leqslant 4\,[1-\delta(\gamma+\varepsilon)]^{2j}\,|f|^2+\gamma\,|\mathring{x}(\varepsilon)|^2\,.$$

Then

$$\begin{gathered}|Ax_j(\varepsilon)-f|=|A\left(x_j(\varepsilon)-\mathring{x}(\varepsilon)\right)+A\mathring{x}(\varepsilon)-f|\leqslant[|A\left(x_j(\varepsilon)-\mathring{x}(\varepsilon)\right)|+|A\mathring{x}(\varepsilon)-f|]\leqslant \\ \leqslant 2|f|\,(1-\delta(\gamma+\varepsilon))^j+|\mathring{x}(\varepsilon)|\,\sqrt{\gamma}+|A\mathring{x}(\varepsilon)-f|\,.\end{gathered}$$

Since

$$(1-\delta(\gamma+\varepsilon))^{2j}\leqslant\gamma$$

we have

$$|Ax_j(\varepsilon)-f|\leqslant 2|f|\sqrt{\gamma}+\sqrt{\gamma}\,|\mathring{x}(\varepsilon)|+|A\mathring{x}(\varepsilon)-f|\,.$$

It implies item a) of the theorem.

Furthermore, using (8) we have

$$\begin{aligned} &|A^*A\,(x_j(\varepsilon)-\mathring{x}(\varepsilon))|^2=\sum_{k=1}^{n} s_k^4\left(1-\delta(s_k^2+\varepsilon)\right)^{2j}|\mathring{x}_k(\varepsilon)|^2= \\ &=\sum_{s_k^2>\gamma} s_k^4\left(1-\delta(s_k^2+\varepsilon)\right)^{2j}|\mathring{x}_k(\varepsilon)|^2+\sum_{s_k^2\leqslant\gamma} s_k^4\left(1-\delta(s_k^2+\varepsilon)\right)^{2j}|\mathring{x}_k(\varepsilon)|^2\leqslant \\ &\leqslant\sum_{s_k^2>\gamma} s_k^4\left(1-\delta(s_k^2+\varepsilon)\right)^{2j}|\mathring{x}_k(\varepsilon)|^2+\sum_{s_k^2\leqslant\gamma} s_k^4\left(1-\delta(s_k^2+\varepsilon)\right)^{2j}|\mathring{x}_k(\varepsilon)|^2 . \end{aligned} \tag{14}$$

It follows

$$|A^*A\,(x_j(\varepsilon)-\mathring{x}(\varepsilon))|^2\leqslant(1-\delta(\gamma+\varepsilon))^{2j}\,|A^*A\mathring{x}|^2+\gamma|\mathring{x}|^2.$$

If j is taken by

$$(1-\delta(\gamma+\varepsilon))^{2j}\leqslant\gamma^2,$$

we obtain

$$|A^*A\,(x_j(\varepsilon)-\mathring{x}(\varepsilon))|^2\leqslant\gamma^2\left[|A^*A\mathring{x}|^2+|\mathring{x}|^2\right].$$

It implies assertion of item b) of the theorem.

By (14) we have

$$\begin{aligned} |A^*A\,(x_j(\varepsilon)-\mathring{x}(\varepsilon))|^2 &\leqslant \sum_{k=1}^{n}(1-\delta(\gamma+\varepsilon))^{2j}\left|s_k^2\mathring{x}_k(\varepsilon)\right|^2+\sum_{k=1}^{n} s_k^2\,|\mathring{x}_k(\varepsilon)|^2= \\ &= (1-\delta(\gamma+\varepsilon))^{2j}\,|A^*A\mathring{x}_k(\varepsilon)|^2+\gamma\,|A^*A\mathring{x}_k(\varepsilon)|^2= \\ &= (1-\delta(\gamma+\varepsilon))^{2j}\,|A^*(A\mathring{x}_k(\varepsilon)-f)+A^*f|^2+\gamma\,|A\mathring{x}_k(\varepsilon)|^2\leqslant \\ &\leqslant 2\,(1-\delta(\gamma+\varepsilon))^{2j}\left(|A^*(A\mathring{x}_k(\varepsilon)-f)|^2+|A^*f|^2\right)+\gamma\,|A\mathring{x}_k(\varepsilon)|^2 . \end{aligned}$$

By applying (13) we get the following inequalities

$$\begin{aligned} |A^*(A\mathring{x}_k-f)|^2 &\leqslant \|A^*\|^2\left(|A\mathring{x}_k|^2+|f|^2\right)\leqslant 5\,\|A^*\|^2\,|f|^2 , \\ |A\mathring{x}_k(\varepsilon)|^2 &\leqslant 4\,|f|^2 . \end{aligned}$$

Therefore

$$\begin{aligned} |A^*A\,(x_j(\varepsilon)-\mathring{x}(\varepsilon))|^2 &\leqslant 2\,(1-\delta(\gamma+\varepsilon))^{2j}\left[5\,\|A^*\|^2\,|f|^2+|A^*f|^2\right]+ \\ +4\gamma\,|f|^2 &\leqslant 10\,(1-\delta(\gamma+\varepsilon))^{2j}\,\|A^*\|^2\,|f|^2+4\gamma\,|f|^2 . \end{aligned}$$

Choosing j from

$$(1-\delta(\gamma+\varepsilon))^{2j}\,10\,\|A^*\|^2\leqslant 4\gamma,$$

we get

$$|A^*A\,(x_j(\varepsilon)-\mathring{x}(\varepsilon))|^2\leqslant 8\gamma\,|f|^2 .$$

This completes the proof of item c) of the theorem.

For $\varepsilon=0$ using Lemma 2.2 we have

$$A^*A\mathring{x}=Af.$$

Therefore item d) of the theorem is proved.

Usually it is important in practice to reduce the difference in $Ax-f$ (less important to find solution of equation $Ax=f$!). Thus, the theorem allows one to solve problem (2) effectively. Note that usage of formula for x_j doesn't require $\varepsilon>0$. Much more suitable case is $\varepsilon=0$.

Now we can suggest the next numerical algorithm based on the Theorem 4.1. We can form sufficiently effective process of solving problem (2) with ill-conditioned or non-invertible matrix. The algorithm will be distinguished from above one only by these points:

It is chosen $\gamma>0$ (stands for accuracy). The number ε is chosen to be zero. The conditions from item d) Theorem 4.1 are verified after every cycle of iteration. Computation is finished when condition (11) holds.

Implementation of the suggested algorithm is realized in K. Satpayev Kazakh National Technical University.

Some of results of this work have been announced in [6] (see also [7]).

References

[1] Baldybek, J., Otelbaev, M., (2001), Parallelization problem of the linear algebraic systems, Mathematician Journal, Almaty, 11(39), pp.53-58, (in Russian).

[2] Bertsekas, Dimitri P., Tsitsiklis, John N., (1997), Parallel and Distributed Computation: Numerical Methods.

[3] Brucel, P.I., (1993), Parallel Programming an Introduction, Prentice Hall International (UK) Limited, 270p.

[4] Gohberg, Israel, Goldberg, Seymour, Marinus A.Kashoek, (2003), Basic Classes of Linear Operators, Birkhauser Verlag, 238p.

[5] Ortega, J.M., (1982), Introduction to Parallel and Vector Solution of Linear Systems, Plenum Press, New York.

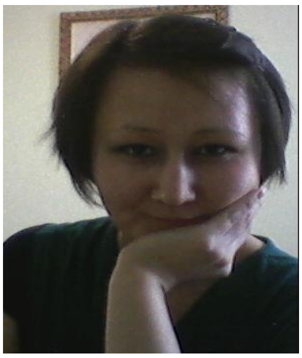
[6] Otelbaev, M., Tuleuov, B., Zhusupova, D., (2011), On a method of finding approximate solutions of ill-conditioned algebraic systems and parallel computation, Eurasian Mathematical Journal, 2(1), pp.149-151.

[7] Otelbaev, M., Zhusupova, D, Tueluov, B., (2011), Parallelization of linear algebraic system with the invertible matrix, Vestnik Bashkirsk. Univ.,16(4), pp. 1129-1133 (in Russian).

[8] Voevodin, V.V., Voevodin, Vl.V., (2002), Paralell Computing, BHV-Peterburg, 608p. (in Russian).

Mukhtarbai Otelbaev was born in 1942 in the Zhambyl region of Kazakhstan. He is a Doctor of Physical-Mathematical Sciences, Professor of Gumilyov Eurasian National University, deputy director of Kazakhstan Branch of Lomonosov Moscow State University. In 1969 he graduated from Mechanical-Mathematical Faculty of Lomonosov Moscow State University. M. Otelbaev got his Ph.D. degree in 1972 and Doctor of Sciences degree in 1978 in Moscow. In 1989 M. Otelbaev was elected a correspondent member and in 2004 an active member of the National Academy of Sciences of the Republic of Kazakhstan.

His research interests are the spectral theory of operators, theory of contraction and expansion operators, attachment theory of functional spaces, approximation theory, computational mathematics and inverse problems. He is an author of more than 200 scientific works, 3 monographers.



Zhusupova Dinara graduated from L.N. Gumilyov Eurasian National University in 2006. She received Ph.D. degree in L.N.Gumilyov Eurasian National University. Her research interests are spectral theory of operators, theory of contraction and expansion operators, computational mathematics.

Berik Tuleuov was born in 1964 in Karaganda province of Kazakhstan. He graduated from Novosibirsk State University in 1988. He has published 15 scientific works.