INSTITUTUL DE MATEMATICA
AL ACADEMIEI ROMANE

PREPRINT SERIES OF THE INSTITUTE OF MATHEMATICS
OF THE ROMANIAN ACADEMY

ISSN 0250 3638

PROJECTIONS AND APPROXIMATE ORTHOGONALIZATION FOR NUMERICAL SOLUTION OF INCONSISTENT LEAST-SQUARES PROBLEMS

by

CONSTANTIN POPA

Preprint nr. 6/1999

BUCURESTI

PROJECTIONS AND APPROXIMATE ORTHOGONALIZATION FOR NUMERICAL SOLUTION OF INCONSISTENT LEAST–SQUARES PROBLEMS

by

CONSTANTIN POPA *

May, 1999

* Faculty of Mathematics and Informatics, "OVIDIUS" University Constanta, Romania.
E–mail: cpopa@ovidius.ct.ro

Projections and approximate orthogonalization for numerical solution of inconsistent least-squares problems

Constantin Popa, Faculty of Mathematics and Informatics, "OVIDIUS" University Constanta , Romania; email : cpopa @ ovidius.ct.ro

Abstract. In a previous paper we described an iterative algorithm for numerical solution of consistent linear least-squares problems. In the present one we generalize it to the case of inconsistent problems. This new algorithm is based on an extension of the classical Kaczmarz's projections method (also obtained by the author in a previous work) and an approximate orthogonalization technique due to Z. Kovarik. We prove that the new algorithm converges to any solution of an inconsistent and rank-defficient least-squares problem (with respect to the choice of the initial approximation).
AMS Subject Classification : 65F10 , 65F20.
Key words and phrases : Kaczmarz's iteration, approximate orthogonalization, inconsistent and rank-defficient least-squares problems.

1 Preliminaries

Let A be a real $m \times n$ matrix and $b \in I\!R^m$. We shall denote by $A^t, (A)_i, (A)^j$, $r(A), R(A), N(A), b_i$ the transpose, i-th row, j-th column, rank, range, null space of A and i-th component of b, respectively (all the vectors that appear being considered as column vectors). The notations $\rho(B), \tau(B)$ will be used for the spectral radius and spectrum of a (square) matrix B and $\| A \|$ will be the spectral norm of A defined by $\| A \|^2 = \rho(A^tA) = \rho(AA^t)$. $P_S(x)$ will be the orthogonal projection of x onto the vector subspace S with respect to the Euclidean scalar product and the associated norm, denoted by $< \cdot, \cdot >$ and $\| \cdot \|$, respectively. We shall consider the linear least-squares problem : find $x^* \in I\!R^n$ such that

$$\| Ax^* - b \| = min! \tag{1}$$

It is well known (see e.g. [1]) that the set of all (least-squares) solutions of (1), denoted by $LSS(A; b)$ is a nonempty closed convex subset of $I\!R^n$ containing a unique solution with minimal norm, denoted by x_{LS}. More than that, we have

$$x^* \in LSS(A;b) \Leftrightarrow A^tAx^* = A^tb \tag{2}$$

and if

$$b_A = P_{R(A)}(b), \tag{3}$$

then

$$LSS(A;b) = S(A;b_A), \tag{4}$$

where by $S(A;b_A)$ we denoted the set of all (classical) solutions of the (consistent) system

$$Ax = b_A. \tag{5}$$

We shall also suppose that the rows and columns of A satisfy

$$(A)_i \neq 0,\ i = 1,\ldots,m;\ \ (A)^j \neq 0,\ j = 1,\ldots,n. \tag{6}$$

Then, we can define the linear applications (matrices)

$$f_i(A;b;x) = x - \frac{< x,(A)_i > - b_i}{\| (A)_i \|^2}(A)_i,\ \varphi_j(A;y) = y - \frac{< y,(A)^j >}{\| (A)^j \|^2}(A)^j, \tag{7}$$

$$K(A;b;x) = (f_1 \circ \ldots \circ f_m)(A;b;x),\ \Phi(A;y) = (\varphi_1 \circ \ldots \circ \varphi_n)(A;y), \tag{8}$$

$$P_i(x) = x - \frac{< x,(A)_i >}{\| (A)_i \|^2}(A)_i,\ \ Q = P_1 \ldots P_m \tag{9}$$

and R the real $n \times m$ matrix of which i-th column $(R)^i$ is given by

$$(R)^i = \frac{1}{\| (A)_i \|^2} P_1 P_2 \ldots P_{i-1}((A)_i) \tag{10}$$

with $P_0 = I$ (the unit matrix). The following results are proved in [8].

Proposition 1 *(i) We have*

$$K(A;b;x) = Qx + Rb,\ Q + RA = I,\ Ry \in R(A^t),\ \forall\ y \in I\!R^m. \tag{11}$$

(ii) $N(A)$ and $R(A^t)$ are invariant subspaces for Q and

$$Q = P_{N(A)} \oplus \tilde{Q},\ \ P_{N(A)}\tilde{Q} = \tilde{Q}P_{N(A)} = 0 \tag{12}$$

where $\tilde{Q}$ is the linear application defined by

$$\tilde{Q} = QP_{R(A^t)}. \tag{13}$$

(iii) $N(A^t)$ and $R(A)$ are invariant subspaces for $\Phi = \Phi(A;\cdot)$ and

$$\Phi = P_{N(A^t)} \oplus \tilde{\Phi}, \quad P_{N(A^t)}\tilde{\Phi} = \tilde{\Phi}P_{N(A^t)} = 0, \tag{14}$$

where $\tilde{\Phi} = \tilde{\Phi}(A;\cdot)$ is the linear application defined by

$$\tilde{\Phi} = \Phi P_{R(A)}. \tag{15}$$

(iv) The applications $\tilde{Q}$ and $\tilde{\Phi}$ satisfy

$$\| \tilde{Q} < 1, \| \tilde{\Phi} \| < 1. \tag{16}$$

The following algorithm (described by the author in [3] and [4]) is an extension of the classical Kaczmarz's projections method (see [8]) : let $x^0 \in I\!R^n, y^0 = b$; for $k = 0, 1, \ldots$ compute

$$y^{k+1} = \Phi(A; y^k), \tag{17}$$

$$\beta^{k+1} = b - y^{k+1}, \tag{18}$$

$$x^{k+1} = K(A; \beta^{k+1}; x^k). \tag{19}$$

In [4] the following are proved.

Proposition 2 *(i) Let G be the $n \times m$ matrix defined by*

$$G = (I - \tilde{Q})^{-1}R. \tag{20}$$

Then, for any matrix A satisfying (6), any $b \in I\!R^m$ and $x^0 \in I\!R^n$ the sequence $(x^k)_{k\geq 0}$ generated with the algorithm (17) - (19) converges and

$$\lim_{k\to\infty} x^k = P_{N(A)}(x^0) + Gb_A. \tag{21}$$

(ii) We have the equalities

$$LSS(A;b) = \{P_{N(A)}(x^0) + Gb_A, \ x^0 \in I\!R^n\}; \quad x_{LS} = Gb_A. \tag{22}$$

Remark 1 *Because the above steps (17) and (19) consist on succesive (orthogonal) projections onto the hyperplanes generated by the columns and rows of A, faster will be the convergence in these steps (and thus for the whole algorithm) if the angles between succesive columns and rows will be closer to 90 degrees (see e.g. [8]).*

Let now $(a_k)_{k\geq 0}$ be the sequence of positive real numbers

$$a_k = \frac{1}{2^{2k}}\frac{(2k)!}{(k!)^2}, \; k \geq 0 \tag{23}$$

and $(q_k)_{k\geq 1}$ a bounded sequence of positive integers, i.e.

$$q_k \leq N, \quad \forall\, k \geq 1. \tag{24}$$

The following algorithms were proposed by Z. Kovarik in [2].
Algorithm (A) Start with $A_0 = A$ and recursively define the matrices H_k and A_{k+1} by

$$H_k = I - A_k A_k^t, \quad \Gamma_k = I + a_1 H_k + \ldots + a_{q_k} H_k^{q_k}, \tag{25}$$

$$A_{k+1} = \Gamma_k A_k, \quad k \geq 0 \tag{26}$$

Algorithm (B) Start with $A_0 = A$ and recursively define the matrices K_k and A_{k+1} by

$$K_k = 2(I + A_k A_k^t)^{-1} - I, \quad \Gamma_k = I + K_k \tag{27}$$

and A_{k+1} as in (26), with Γ_k from (27).
Let A_∞ be the $m \times n$ matrix defined by

$$A_\infty = [(AA^t)^{\frac{1}{2}}]^+ A, \tag{28}$$

where by B^+ we denoted the Moore-Penrose pseudoinverse of the matrix B (see e.g. [1]). We shall also suppose that

$$\| AA^t \| = \| A^t A \| = \rho(AA^t) < 1. \tag{29}$$

In [5] and [7] we applied the above algorithms (analysed in [2] for matrices with linearly independent rows) to an arbitrary rectangular matrix A. The following results were proved.

Proposition 3 *(i) If (24) and (29) hold, then the sequence of matrices $(A_k)_{k\geq 0}$ generated with the algorithm (25) - (26) converges to A_∞.*
(ii) If (29) holds, then the sequence of matrices $(A_k)_{k\geq 0}$ generated with the algorithm (27) converges to A_∞.

Remark 2 *If A has linearly independent rows then we can replace the pseudoinverse $[(AA^t)^{\frac{1}{2}}]^+$ in (28) by the classical inverse $[(AA^t)^{\frac{1}{2}}]^{-1}$ and it can be proved that A_∞ has in this case mutually orthogonal rows (see [2]). This is* *no longer true for a general rectangular A, but an improvement is obtained concerning the angles between succesive rows of A by comparing them with the angles between the rows of the initial matrix A;(see e.g. the numerical experiments from [5] and [7]).*

Remark 3 *The condition (29) is not restrictive. It can be fullfiled by an appropriate scalling of the elements of the matrix A.*

2 Auxiliary results

In order to "mix" the above Kaczmarz and Kovarik algorithms and to prove convergence of the new one so obtained, we need some preparatory results which will be presented in this section. First of all we are interested in the fulfilment of assumption (6) for any matrix A_k generated by one of the above Kovarik's algorithms **(A)** or **(B)**. This problem has been already analysed and solved in [6], thus we will only remind here the corresponding results.

Proposition 4 *Let us suppose that (6) holds for A. Then it also holds for any matrix A_k generated with the algorithm (25) - (26) or the algorithm (27), i.e.*

$$(A_k)_i \neq 0,\ i = 1, \ldots, m;\ \ (A_k)^j \neq 0,\ j = 1, \ldots, n,\ \forall\ k \geq 0. \tag{30}$$

Let now F_k, L_k be the matrices defined by (see (25), (27))

$$F_k = I - A_k^t A_k,\quad L_k = 2(I + A_k^t A_k)^{-1} - I \tag{31}$$

and

$$B_k = I + a_1 F_k + \ldots + a_{q_k} F_k^{q_k} \tag{32}$$

or

$$B_k = I + L_k. \tag{33}$$

The following result gives supplementary informations about the matrices appearing in algorithms **(A)** and **(B)** and the above defined ones.

Proposition 5 *Let us suppose that (29 holds). Then*
(i) the matrices Γ_k, from (25) or (27), and B_k, from (32) or (33) are symmetric and positive definite ((SPD), for short);
(ii) the following equalities hold

$$\Gamma_k(A_k A_k^t) = (A_k A_k^t)\Gamma_k;\ \ B_k(A_k^t A_k) = (A_k^t A_k)B_k, \tag{34}$$

$$A_{k+1} = \Gamma_k A_k = A_k B_k, \tag{35}$$

$$A_k^t \Gamma_k = B_k A_k^t, \tag{36}$$

$$A_{k+1}^t = B_k A_k^t. \tag{37}$$

Proof. (i) In [5] and [7] we proved that, under the assumption (29), the matrices H_k and K_k from (25) and (27), respectively, are (SPD) and

$$\tau(H_k) \subset (0,1], \quad \tau(K_k) \subset (0,1], \quad \forall\, k \geq 0. \tag{38}$$

From (38), (23), (25) and (27) it results that Γ_k is (SPD). Similar arguments tell us that F_k and L_k from (31) are (SPD),

$$\tau(F_k) \subset (0,1], \quad \tau(L_k) \subset (0,1] \tag{39}$$

and B_k from (32) or (33) is also (SPD).
(ii) If Γ_k, B_k are defined as in (25) and (32), respectively then the equalities in (34) obviously hold. Let now Γ_k be defined as in (27). Then, the first equality in (34) holds from the following sequences of equivalences (the last one being obviously true)

$$K_k(A_k {A_k}^t) = (A_k A_k^t)K_k \Leftrightarrow (I + A_k A_k^t)K_k(A_k A_k^t)(I + A_k A_k^t) =$$
$$= (I + A_k A_k^t)(A_k A_k^t)K_k(I + A_k A_k^t) \Leftrightarrow$$
$$\Leftrightarrow (I - A_k A_k^t)(A_k A_k^t)(I + A_k A_k^t) = (I + A_k A_k^t)(A_k A_k^t)(I - A_k A_k^t) \Leftrightarrow$$
$$\Leftrightarrow [A_k A_k^t - (A_k A_k^t)^2](I + A_k A_k^t) = [A_k A_k^t + (A_k A_k^t)^2](I - A_k A_k^t) \Leftrightarrow$$
$$A_k A_k^t - (A_k A_k^t)^3 = A_k A_k^t - (A_k A_k^t)^3. \tag{40}$$

Using similar arguments we can prove that

$$L_k(A_k^t A_k) = (A_k^t A_k)L_k \tag{41}$$

which ensures the second equality in (34) for B_k defined as in (33). We can easily observe that (36) and (37) hold from (35) and (26). If Γ_k , B_k are defined as in (25) and (32), then (35) easily holds by observing that

$$H_k A_k = (I - A_k A_k^t)A_k = A_k(I - A_k^t A_k) = A_k F_k. \tag{42}$$

Let now Γ_k, B_k be given by (27) and (33), respectively. We have the following sequence of equivalences

$$K_k A_k = A_k L_k \Leftrightarrow (I+A_k A_k^t)K_k A_k(I+A_k^t A_k) = (I+A_k A_k^t)A_k L_k(I+A_k^t A_k) \Leftrightarrow$$

$$(I - A_kA_k^t)A_k(I + A_k^tA_k) = (I + A_kA_k^t)A_k(I - A_k^tA_k) \Leftrightarrow$$
$$A_k - A_kA_k^tA_kA_k^tA_k = A_k - A_kA_k^tA_kA_k^tA_k \quad (43)$$

Because the last equality is obviously true, from (43), (27) and (33) we obtain (35) in this case too and the proof is complete.

Proposition 6 *If (29) holds and $(A_k)_{k\geq 0}$ is the sequence of matrices defined with the above algorithms* **(A)** *or* **(B)**, *then*

$$N(A_k) = N(A), \quad N(A_k^t) = N(A^t), \quad (44)$$

thus

$$P_{N(A_k)} = P_{N(A)}, \quad P_{N(A_k^t)} = P_{N(A^t)}, \quad \forall\ k \geq 0. \quad (45)$$

Proof. From (26) and (37) it directly holds that

$$N(A_k) \subset N(A_{k+1}), \quad N(A_k^t) \subset N(A_{k+1}^t), \quad \forall\, k \geq 0. \quad (46)$$

Let $z \in N(A_{k+1})$. Then, from (26) and the fact that Γ_k is (SPD) (thus invertible) we obtain that $A_kz = 0$, i.e. $z \in N(A_k)$, thus

$$N(A_{k+1}) \subset N(A_k), \quad \forall\, k \geq 0.$$

In a similar way, using (37) we obtain that

$$N(A_{k+1}^t) \subset N(A_k^t), \quad \forall\, k \geq 0,$$

which together with (46) gives us (44) (thus (45)) and completes the proof.

3 The Kaczmarz-Kovarik algorithm

Using the constructions and results from the previous sections we can define our new algorithm as follows.
Kaczmarz-Kovarik algorithm: let $x^0 \in I\!R^n$, $A_0 = A$, $b^0 = b$ and

$$H_0 = I - A_0A_0^t \ \ or \ \ K_0 = 2(I + A_0A_0^t)^{-1} - I. \quad (47)$$

Step 1. Compute A_{k+1} and b^{k+1} by

$$A_{k+1} = \Gamma_kA_k, \quad b^{k+1} = \Gamma_kb^k, \quad (48)$$

with Γ_k from (25) or Γ_k from (27), respectively.
Step 2. Compute y^{k+1} and β^{k+1} by

$$y^{k+1} = \Phi^{k+1}(A_{k+1}; b^{k+1}), \tag{49}$$

$$\beta^{k+1} = b^{k+1} - y^{k+1}. \tag{50}$$

Step 3. Compute the next approximation x^{k+1} by

$$x^{k+1} = K(A_{k+1}; \beta^{k+1}; x^k) \tag{51}$$

and update H_k or K_k to H_{k+1} or K_{k+1} by

$$H_{k+1} = I - A_{k+1}A_{k+1}^t \quad or \quad K_{k+1} = 2(I + A_{k+1}A_{k+1}^t)^{-1} - I. \tag{52}$$

Remark 4 *From Proposition 4 it results that the above steps (49) and (51) are well defined for any* $k \geq 0$.

Remark 5 *The step (49) means the succesive application of* $\Phi(A_{k+1}; \cdot)$ $(k+1)$ *- times to the initial vector* b^{k+1}, *i.e.*

$$\Phi^{k+1}(A_{k+1}; b^{k+1}) = (\Phi(A_{k+1}; \cdot) \circ \ldots \circ \Phi(A_{k+1}; \cdot))(b^{k+1}). \tag{53}$$

Remark 6 *In fact, we defined two* **Kaczmarz-Kovarik** *algorithms which corresponds to the two versions* **(A)** *and* **(B)** *of Kovarik's method: the first one starts with* H_0 *from(47), makes the computations in (48) (and after that those in (49)-(51)) with* Γ_k *from (25) and then updates* H_k *to* H_{k+1} *as in (52); the second one starts with* K_0 *from (47), then uses* Γ_k *from (27) in (48)-(51) and updates* K_k *to* K_{k+1} *as in (52).*

The following result ensures us that the set $LSS(A; b)$, corresponding to the initial problem (1) does not change during the transformations (48).

Proposition 7 *We have*

$$LSS(A_k; b^k) = LSS(A; b), \ \forall\ k \geq 0. \tag{54}$$

Proof. Using (2), (37), (34), (36) and the fact that the matrix B_k from (32) is invertible, we can write the following sequence of equivalencies

$$x \in LSS(A_{k+1}; b^{k+1}) \Leftrightarrow A_{k+1}^t A_{k+1} x = A_{k+1}^t b^{k+1} \Leftrightarrow$$

$$B_k A_k^t A_k B_k x = B_k A_k^t \Gamma_k b^k \Leftrightarrow (B_k)^2 A_k^t A_k x = (B_k)^2 A_k^t b^k \Leftrightarrow$$

$$A_k^t A_k x = A_k^t b^k \Leftrightarrow x \in LSS(A_k; b^k),\ \forall\ k \geq 0. \tag{55}$$

From (55) and the fact that $A_0 = A, b^0 = b$, the equalities (54) directly result and the proof is complete.
Let now $Q_k, \tilde{Q}_k, \Phi_k, \tilde{\Phi}_k, R_k$ and G_k be the matrices defined as in (9), (13), (8), (15), (10) and (20), respectively, but with A_k from (48) instead of A. Then, as in the proof of Theorem 1 from [6] we obtain.

Proposition 8 *If $\tilde{Q}_\infty, \tilde{\Phi}_\infty$ and R_∞ are the matrices defined as in (13), (15) and (10), respectively, but with A_∞ from (28) instead of A, then*

$$\lim_{k\to\infty} \tilde{Q}_k = \tilde{Q}_\infty,\ \lim_{k\to\infty} \tilde{\Phi}_k = \tilde{\Phi}_\infty,\ \lim_{k\to\infty} R_k = R_\infty. \tag{56}$$

Following the ideas from [4] we can prove

Proposition 9 *Let $(x^k)_{k\geq 0}$ be the sequence generated with the algorithm (47)-(52). Then*

$$P_{N(A_k)}(x^k) = P_{N(A)}(x^k) = P_{N(A)}(x^0),\ \forall\ k \geq 0. \tag{57}$$

Proof. We shall use the mathematical induction. For $k = 0$, (57) is true. Then, let $k \geq 0$ be fixed. By using (11), (51) and (12) we obtain

$$x^{k+1} = Q_{k+1}x^k + R_{k+1}\beta^{k+1} = P_{N(A_{k+1})}(x^k) + \tilde{Q}_{k+1}(x^k) + R_{k+1}\beta^k. \tag{58}$$

But, from (11) with A_{k+1} and R_{k+1} instead of A and R, respectively, and (13) with A_{k+1} and $Q_{k+1}, \tilde{Q}_{k+1}$ instead of A, Q and $\tilde{Q}$, we obtain

$$R_{k+1}\beta^k \in R(A_{k+1}^t),\ \tilde{Q}_{k+1}(x^k) \in R(A_{k+1}^t). \tag{59}$$

From (58) and (59) we obtain $P_{N(A_{k+1})}(x^{k+1}) = P_{N(A_{k+1})}(x^k)$. The remaining equalities in (57) then hold from (44)-(45) and the proof is complete.
The last preparatory result before our main theorem is the following proposition, concerning the sequence of "right hand sides" $(b^k)_{k\geq 0}$ from (48).

Proposition 10 *Let $(b^k)_{k\geq 0}$ be as in (48) and, for any $k \geq 0$*

$$b_{A_k}^k = P_{R(A_k)}(b^k). \tag{60}$$

Then, the sequence $(b_{A_k}^k)_{k\geq 0}$ is bounded.

Proof. Let us suppose that the conclusion of our proposition is false. Then, it would exist a subsequence of $(b^k_{A_k})_{k\geq 0}$ (which, for simplicity we shall denote in the same way) such that

$$\lim_{k\to\infty} \| b^k_{A_k} \| = +\infty. \tag{61}$$

From (5) and (60) it holds that

$$x \in LSS(A_k : b^k) \Leftrightarrow A_k x = b^k_{A_k}. \tag{62}$$

Let now $x^* \in LSS(A; b)$. From (54) and (62) we obtain

$$A_k x^* = b^k_{A_k}, \quad \forall \ k \geq 0. \tag{63}$$

From Proposition 3 we have $\lim_{k\to\infty} A_k = A_\infty$, which tells us that it exists an integer $k_0 \geq 1$ such that

$$\| A_k x^* \| \leq \| A_\infty x^* \| + 1. \tag{64}$$

Let $k_1 \geq k_0 \geq 1$ be such that (see (61))

$$\| b^k_{A_k} \| > \| A_\infty x^* \| + 1, \ \forall \ k \geq k_1. \tag{65}$$

From (63)-(65) we then get a contradiction which completes our proof.
We are now able to prove the main result of the paper, concerning the convergence of the sequence $(x^k)_{k\geq 0}$ generated with the above **Kaczmarz-Kovarik** algorithm.

Theorem 1 *For any $x^0 \in I\!R^n$ if $(x^k)_{k\geq 0}$ is the sequence generated with the algorithm (47)-(52), then*

$$\lim_{k\to\infty} x^k = P_{N(A)}(x^0) + G b_A. \tag{66}$$

Proof. Let $k \geq 0$ be arbitrary fixed and $b^k_* \in I\!R^m$ defined by

$$b^k_* = P_{N(A^t_k)}(b^k). \tag{67}$$

Then, we have the orthogonal decomposition of b^k (see (60))

$$b^k = b^k_{A_k} \oplus b^k_* \tag{68}$$

and, from [4] (Theorem 1 and Proposition 1)

$$LSS(A_k; b^k) = \{P_{N(A_k)}(x^0) + G_k b^k_{A_k}, x^0 \in I\!R^n\}, \tag{69}$$

$$x_{LS} = G_k b^k_{A_k} = G b_A. \tag{70}$$

Let now $x^0 \in I\!R^n$ be arbitrary fixed. Then, from (70), (45), (11), (51), (12) and (57) we succesively get

$$x^{k+1} - (P_{N(A)}(x^0) + Gb_A) = x^{k+1} - (P_{N(A_{k+1})}(x^0) + G_{k+1} b^{k+1}_{A_{k+1}}) =$$

$$(P_{N(A_{k+1})}(x^k) + \tilde{Q}_{k+1} x^k + R_{k+1}\beta^{k+1}) - (P_{N(A_{k+1})}(x^k) + G_{k+1} b^{k+1}_{A_{k+1}}) =$$

$$\tilde{Q}_{k+1} x^k + R_{k+1}\beta^{k+1} - [(I - \tilde{Q}_{k+1}) + \tilde{Q}_{k+1}][(I - \tilde{Q}_{k+1})^{-1} R_{k+1}] b^{k+1}_{A_{k+1}} =$$

$$\tilde{Q}_{k+1} x^k + R_{k+1}\beta^{k+1} - R_{k+1} b^{k+1}_{A_{k+1}} - \tilde{Q}_{k+1} G_{k+1} b^{k+1}_{A_{k+1}} - \tilde{Q}_{k+1} P_{N(A_{k+1})}(x^0) =$$

$$\tilde{Q}_{k+1}[x^k - (P_{N(A)}(x^0) + Gb_A)] + R_{k+1}(\beta^{k+1} - b^{k+1}_{A_{k+1}}). \tag{71}$$

Now, from (50), (68), (49), (14) and (67) we obtain

$$\beta^{k+1} - b^{k+1}_{A_{k+1}} = b^{k+1} - y^{k+1} - b^{k+1}_{A_{k+1}} = b^{k+1}_* - y^{k+1} = b^{k+1}_* - \Phi^{k+1}(A_{k+1}; b^{k+1}) =$$

$$b^{k+1}_* - [P_{N(A^t_{k+1})} \oplus \tilde{\Phi}_{k+1}]^{k+1}(b^{k+1}) = b^{k+1}_* - [P_{N(A^t_{k+1})} \oplus (\tilde{\Phi}_{k+1})^{k+1}](b^{k+1}) =$$

$$[b^{k+1}_* - P_{N(A^t_{k+1})}(b^{k+1})] - (\tilde{\Phi}_{k+1})^{k+1}(b^{k+1}) =$$

$$-(\tilde{\Phi}_{k+1})^{k+1}(b^{k+1}) = -(\tilde{\Phi}_{k+1})^{k+1}(b^{k+1}_{A_{k+1}}) \tag{72}$$

Let $x^* \in I\!R^n$ be defined by (see (66))

$$x^* = P_{N(A)}(x^0) + Gb_A. \tag{73}$$

Then, from (71) and (72) we obtain

$$x^{k+1} - x^* = \tilde{Q}_{k+1}(x^k - x^*) - R_{k+1}(\tilde{\Phi}_{k+1})^{k+1}(b^{k+1}_{A_{k+1}}), \ \forall \ k \geq 0. \tag{74}$$

By iterating the equality (74) we get

$$x^{k+1} - x^* = \tilde{Q}_{k+1} \dots \tilde{Q}_1(x^0 - x^*) -$$

$$\sum_{j=1}^{k} \tilde{Q}_{k+1} \dots \tilde{Q}_{j+1} R_j (\tilde{\Phi}_j)^j (b^j_{A_j}) - R_{k+1}(\tilde{\Phi}_{k+1})^{k+1}(b^{k+1}_{A_{k+1}}) \tag{75}$$

Then, by taking norms, from (75) it results

$$\| x^{k+1} - x^* \| \leq \| \tilde{Q}_{k+1} \| \dots \| \tilde{Q}_1 \| \| x^0 - x^* \| +$$

$$\sum_{j=1}^{k}(\| \tilde{Q}_{k+1} \| \dots \| \tilde{Q}_{j+1} \| \| \| \tilde{\Phi}_j \|^j \| R_j \| \| \| b^j_{A_j} \|) + \| R_{k+1} \| \| \| \tilde{\Phi}_{k+1} \|^{k+1} \| b^{k+1}_{A_{k+1}} \| . \tag{76}$$

From (16) we obtain that

$$\| \tilde{Q}_k \| < 1, \ \| \tilde{\Phi}_k \| < 1, \ \forall\, k \geq 0, \ \| \tilde{Q}_\infty \| < 1, \ \| \tilde{\Phi}_\infty \| < 1. \tag{77}$$

Let then $k_0 \geq 1$ and $M_0 > 0$ be such that

$$\| \tilde{Q}_k \| < \frac{1 + \| \tilde{Q}_\infty \|}{2} < 1, \quad \| \tilde{\Phi}_k \| < \frac{1 + \| \tilde{\Phi}_\infty \|}{2} < 1, \tag{78}$$

$$\| R_k \| < \| R_\infty \| + 1, \quad \| b^{k+1}_{A_{k+1}} \| \leq M_0, \ \forall\, k > k_0 \tag{79}$$

(such k_0 and M_0 exist according to Proposition 10 and (56)). Let now $\rho, \mu \in (0,1)$ and $M > 0$ be defined by

$$\rho = \max\{\| \tilde{Q}_1 \|, \dots, \| \tilde{Q}_{k_0} \|, \frac{1 + \| \tilde{Q}_\infty \|}{2}\}, \tag{80}$$

$$\mu = \max\{\| \tilde{\Phi}_1 \|, \dots, \| \tilde{\Phi}_{k_0} \|, \frac{1 + \| \tilde{\Phi}_\infty \|}{2}\}, \tag{81}$$

$$M = \max\{\| R_1 \|, \dots, \| R_{k_0} \|, \| R_\infty \| + 1, \| b^0_{A_0} \|, \dots, \| b^{k_0}_{A_{k_0}} \|, M_0\}. \tag{82}$$

Then, from (76)-(82) we get

$$\| x^{k+1} - x^* \| \leq \rho^{k+1} \| x^0 - x^* \| + M^2 (\sum_{j=1}^{k} \rho^{k+1-j} \mu^j + \mu^{k+1}), \ \forall\, k \geq 0. \tag{83}$$

Now, by choosing γ as

$$\gamma = \max\{\rho, \mu\} \in (0,1), \tag{84}$$

from (83) we obtain

$$\| x^{k+1} - x^* \| \leq \gamma^{k+1} (\| x^0 - x^* \| + M^2 (k+1)), \ \forall\, k \geq 0. \tag{85}$$

From (84) and (85) we get

$$\lim_{k \to \infty} \| x^{k+1} - x^* \| = 0$$

and the proof is complete.
As in [4] (see also Proposition 2) the following consequence of the above theorem can be proved.

Corollary 1 *In the hypothesis of Theorem 1, for any* $x^0 \in I\!R^n$ *the limit point (66) of the Kaczmarz-Kovarik sequence* $(x^k)_{k \geq 0}$ *is a solution of the problem (1). More than that, its minimal norm solution* x_{LS} *is obtained as a limit point in (66) if and only if* $x^0 \in R(A^t)$.

References

[1] Å. Björck, *Numerical methods for least squares problems*, SIAM Philadelphia, 1996.

[2] Z. Kovarik, *Some iterative methods for improving orthonormality*, SIAM J. Numer. Anal., **7(3)** (1970), pp. 386–389.

[3] C. Popa, *Least-squares solution of overdetermined inconsistent linear systems using Kaczmarz's relaxation*, Intern. J. Comp. Math., **55** (1995), pp. 79–89.

[4] C. Popa, *Characterization of the solution set of inconsistent least-squares problems by an extended Kaczmarz algorithm*, Korean J. Comp. Appl. Math., **6(1)** (1999), pp. 51–64.

[5] C. Popa, *Extension of an approximate orthogonalization algorithm to arbitrary rectangular matrices*, Preprint nr. **18** (1998), Institute of Math. of the Romanian Academy, Bucharest.

[6] C. Popa, *A fast Kaczmarz-Kovarik algorithm for least-squares problems*, submitted to *B I T*, 1999.

[7] C. Popa, *An iterative method for improving orthogonality of rows of arbitrary rectangular matrices*, submitted to **CALCOLO**, 1999.

[8] K. Tanabe, *Projection method for solving a singular system of linear equations and its applications*, Numer. Math., **17** (1971), pp. 203-214.