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

Lehrstuhl für Informatik 10 (Systemsimulation)

A hybrid Kaczmarz - Conjugate Gradient algorithm

Constantin Popa

Lehrstuhlbericht 07-2

A hybrid Kaczmarz - Conjugate Gradient algorithm

Constantin Popa*

Abstract

In this paper we present a hybrid iterative algorithm for solving inconsistent linear least squares problems arising in image reconstruction from projections in computerized tomography. It includes in each iteration a CG-like step for modifying the right hand side and a Kaczmarz step for producing the approximate image. We prove convergence of the hybrid algorithm for inconsistent and rank-deficient problems. Numerical experiments on a regularized image reconstruction problem involving a real 2D medical data set are also presented in the last section of the paper.

2000 MS Classification: 65F10, 65F20

Key words and phrases: extended Kaczmarz algorithm, Conjugate Gradient method, inconsistent rank-deficient least squares problems, image reconstruction from projections, Tikhonov regularization

1 Introduction

The Image Reconstruction from Projections (IRP) is essentially based on the researches and results due to A. M. Cormack and G. Hounsfield in early 50's, which combined their efforts in the construction of the first computer tomograph for medical investigations, in 1972 (see [7], [3], [12]). For modelling the IRP problem several classes of methods have been developed starting with 80's. Among them, the Algebraic Reconstruction Techniques (ART) have a special position. The main and common idea of these methods is to transform the initial (integral equation) formulation of the IRP problem into a linear system of equalities (or inequalities - this case will not be analysed in the present paper), involving a matrix A and a right hand side b. The matrix A contains information related to the "image scanning procedure", whereas the vector b (essentially) comes from measurements of the X-rays intensities (see [7]). Usually A is rectangular ($m \times n$), big, sparse and rank-deficient, and $b \in I\!R^m$ is affected by the measurement errors, all these aspects giving us finally a least squares formulation of the problem as follows: find $x \in I\!R^n$ such that

$$\parallel Ax - b \parallel = \min! \tag{1}$$

($\langle \cdot, \cdot \rangle$ and $\parallel \cdot \parallel$ will denote the euclidean scalar product and norm, respectively). For the numerical solution of (1) a large class of iterative algorithms have been developed. These are the "row-action methods", which use in each iteration rows or blocks of rows from the matrix (see [4, 3, 12]). This aspect is also directly related to the construction of A during the scanning procedure - the i-th X-ray intersects some of the pixels P_j following segments of lenght A_{ij}, which will give us the corresponding (nonzero) elements on the i-th row of A (see Figure 1). Moreover, for very big dimensions m and n (as e.g. in 3D medical applications; see [16]), in which case A cannot be anymore stored (even in a compressed form), this particular property of the "row-action methods" gives us the possibility to (re)generate in each iteration the corresponding rows or blocks of rows from A and use them successively (or simultaneously), as it is ilustrated by the examples from below. For this, let A^T, A_i, A^j be the transpose, i-th row and j-th column of A, respectively; moreover, all the vectors that will appear are column vectors. We shall first refer at computations of the form

*"Ovidius" University of Constanta, Romania; e-mail: cpopa@univ-ovidius.ro. The paper was supported by the project CEEX 05-D11-25/2005

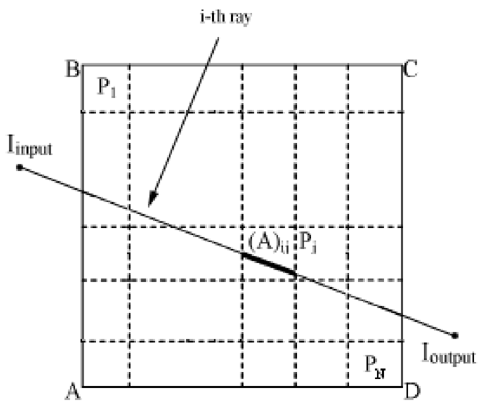

Figure 1: Construction of A and b

$u = (u_1, \ldots, u_m)^T = Ax \Leftrightarrow u_i = \langle x, A_i \rangle, \forall i = 1, \ldots, m$ and $v = A^T y$ (used by many "row-action methods" in their iterations; see [4]). These can be easily obtained by successive row generation of A as follows.

Example 1 : compute $u = Ax$ and $v = A^T y$
$v = 0$
for $i = 1 : m$
generate the $i-$th row A_i
$u_i = \langle x, A_i \rangle$
$v = v + y_i \cdot A_i$
endfor

Unfortunately, the above simple algorithm cannot be used when the situation from the next example appears.

Example 2 : successive computations using the columns $A^j, j = 1, \ldots, n$
for $j = 1 : n$
$aux = \frac{\langle y, A^j \rangle}{\| A^j \|^2}$
$y = y - aux \cdot A^j$
endfor

Indeed, "extracting" a column A^j from a "row-genarated" matrix A is a very time consuming task which in practical applications can determine the "rejection" of the corresponding method. Our paper is motivated and concerned with the above mentioned aspects. In this sense, in section 2 we briefly present two iterative "row-action" solvers for inconsistent problems as (1) - the Kaczmarz Extended algorithm (KE), previously introduced by the author in [13] (see also [14]) and the Generalized Conjugate Gradient method (GCG), see [10] and references therein and also [18] and [6]. The KE algorithm, which generalizes the classical Kaczmarz's one (see [9]) gave good results in both medical and geodesic image reconstruction problems, in classical or regularized formulations (see [1, 15]), whereas GCG seems to be less appropriate for such kind of problems (see [7, 8]). But unfortunately, KE uses in each iteration column computations as in Example 2, whereas in the GCG iteration we are in the case from Example 1. In this sense, in section 3 of the paper we "mix" the above algorithms in a "hybrid" version in which we replace a step as in Example 2, from the KE method with a GCG-like algorithm, which belongs to the class of methods from Example 1. We prove convergence of the new algorithm so obtained, for general inconsistent least squares problems as (1). In section 4 we present applications and comparisons of our mixed algorithm with the KE one for Tikhonov-like regularized formulations of a 2D IRP problem arising in medical applications.

2 The Extended Kaczmarz and Generalized CG algorithms

In the rest of the paper we shall denote by $LSS(A;b), x_{LS}$ the set of all least squares solutions of (1) and the minimal norm one, and by P_S the orthogonal projection onto a subspace $S \subset I\!R^q$.
Algorithm KE. Let $x^0 \in I\!R^n, y^0 = b$; for $k = 0, 1, \ldots$ do

$$\begin{aligned} y^{k+1} &= (\varphi_1 \circ \ldots \circ \varphi_n)(y^k) && (2) \\ b^{k+1} &= b - y^{k+1} && (3) \\ x^{k+1} &= (f_1 \circ \ldots \circ f_m)(b^{k+1}; x^k) && (4) \end{aligned}$$

where

$$\varphi_j(y) = y - \frac{\langle y, A^j \rangle}{\|A^j\|^2} A^j, \quad f_i(\beta; x) = x - \frac{\langle x, A_i \rangle - \beta_i}{\|A_i\|^2} A_i \tag{5}$$

Theorem 1 ([14]) *Let us suppose that $A_i \neq 0, A^j \neq 0, i = 1, \ldots, m, j = 1, \ldots, n$. Then, for any $x^0 \in I\!R^n$ the above sequence $(x^k)_{k \geq 0}$ converges and*

$$\lim_{k \to \infty} x^k = P_{N(A)}(x^0) + x_{LS} \in LSS(A; b). \tag{6}$$

If $x^0 \in R(A^T)$ then $\lim_{k \to \infty} x^k = x_{LS}$.

Remark 1 *The sequence $(y^k)_{k \geq 0}$ constructed in the step (2) of the KE algorithm converges and*

$$\lim_{k \to \infty} y^k = P_{N(A^T)}(b). \tag{7}$$

Moreover, the step (2) represents the classical Kaczmarz algorithm (i.e. step (4) with b instead of b^{k+1}) applied to the (consistent) system

$$A^T y = 0. \tag{8}$$

Algorithm GCG ([10], [18], [6]) Let $x^0 \in I\!R^n, r^0 = b - Ax^0$; for $k = 0, 1, \ldots$ do

$$\begin{aligned} r^0 &= b - Ax^0, \\ p^0 &= A^T r^0, \\ \alpha_k &= \| A^T r^k \|^2 \, / \, \| A p^k \|^2, \\ x^{k+1} &= x^k + \alpha_k p^k, \\ r^{k+1} &= b - Ax^{k+1} = r^k - \alpha_k A p^k, \\ \beta_k &= \| A^T r_{k+1} \|^2 \, / \, \| A^T r^k \|^2, \\ p^{k+1} &= A^T r^{k+1} + \beta_k p^k. && (9) \end{aligned}$$

Theorem 2 ([10]) *Let $r = \mathrm{rank}(A)$, $U^T A V = \mathrm{diag}(\sigma_1, \ldots, \sigma_r, 0, \ldots, 0)$ a singular value decomposition of A, with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$ and the numbers δ and c defined by*

$$\delta = \frac{\sigma_1 - \sigma_r}{\sigma_1 + \sigma_r} \in [0, 1), \quad c = \frac{\| Ax^0 - P_{R(A)}(b) \|}{\sigma_r}. \tag{10}$$

Then, for any $x^0 \in I\!R^n$ the sequence $(x^k)_{k \geq 0}$ generated with the GCG algorithm converges to the same vector from (6) and we have the estimate

$$\| x^k - (P_{N(A)}(x^0) + x_{LS}) \| \leq \ c \cdot \delta^k, \ \forall k \geq 1. \tag{11}$$

If we apply the GCG algorithm to the system (8) with $y^0 = b$, we obtain a sequence $(y^k)_{k \geq 0}$ convergent to $P_{N(A^T)}(b)$. Moreover, from (11) we get the estimate

$$\| y^k - P_{N(A^T)} \| \leq \ \gamma \cdot \delta^k, \ \forall k \geq 1, \tag{12}$$

with δ from (10) and γ given by

$$\gamma = \frac{\| A^T b \|}{\sigma_r}. \tag{13}$$

We shall denote the k-th iteration of this algorithm by $y^k = GCG(A^T; y^{k-1})$.

Remark 2 *Unfortunately, inspite of its good behaviour for systems arising from discretization of partial differential equations together with its fully parallelizable property, the GCG algorithm seems to be not appropriate for problems like* (1) *appearing in IRP (see* [7, 8] *and the experiments presented in figures* 4 *and* 5*).*

3 The mixed Kaczmarz - CG algorithm

According to the type of computations during the KE and GCG iterations, we can easily observe that:

- KE Step (4) and the whole algorithm GCG are concerned with computations as in Example 1;
- KE Step (2) is concerned with computations as in Example 2;

Moreover, in the KE iteration, the step in which the image is reconstructed is essentially (4), whereas steps (2) and (3) produce only a new right-hand side b^{k+1}, with the property (see [17, 14])

$$\lim_{k\to\infty} b^{k+1} = b - P_{N(A^T)}(b) = P_{R(A)}(b). \tag{14}$$

All the above arguments lead us to the construction of the following mixed Kaczmarz - Conjugate Gradient algorithm.
Algorithm KECG. Let $x^0 \in I\!R^n, y^0 = b$; for $k = 0, 1, \ldots$ do

$$y^{k+1} = GCG(A^T; y^k) \tag{15}$$

$$b^{k+1} = b - y^{k+1} \tag{16}$$

$$x^{k+1} = (f_1 \circ \ldots \circ f_m)(b^{k+1}; x^k) \tag{17}$$

For proving convergence of the sequence $(x^k)_{k\geq 0}$ generated with the above algorithm KECG we shall replay some notations and properties related to the KE algorithm (for details see [2, 3, 7, 14, 17]). The orthogonal projections $f_i(\beta; x)$ from (5) can be splitted as

$$f_i(\beta; x) = P_i(x) + \frac{\beta_i}{\parallel A_i \parallel^2} A_i, \quad \text{with} \quad P_i(x) = x - \frac{\langle x, A_i \rangle}{\parallel A_i \parallel^2} A_i. \tag{18}$$

Then we define the matrices

$$Q_0 = I,\ Q_i = P_1 \cdot \ldots \cdot P_i,\ i = 1, \ldots, m-1,\ Q = P_1 \cdot \ldots \cdot P_m \tag{19}$$

and

$$F^i = \frac{1}{\parallel A_i \parallel^2} Q_{i-1} A_i,\ i = 1, \ldots, m, \quad F = \mathrm{col}\left(F^1, F^2, \ldots, F^m\right). \tag{20}$$

The following properties hold

$$(f_1 \circ \ldots \circ f_m)(\beta; x) = Qx + F\beta, \quad Q + FA = I, \tag{21}$$

$$Q = P_{N(A)} \oplus \tilde{Q}, \quad \tilde{Q} = QP_{R(A^T)},\ \parallel \tilde{Q} \parallel < 1, \tag{22}$$

$$\tilde{Q} P_{N(A)} = P_{N(A)} \tilde{Q} = 0,\ Fy \in R(A^T),\ \forall y \in I\!R^m. \tag{23}$$

Moreover, the minimal norm solution of (1), x_{LS} is given by

$$x_{LS} = GP_{R(A)}(b), \quad \text{with} \quad G = (I - \tilde{Q})^{-1} F. \tag{24}$$

Lemma 1 *If $(x^k)_{k\geq 0}$ is the sequence generated with the algorithm KECG, we have*

$$P_{N(A)}(x^k) = P_{N(A)}(x^0),\ \forall k \geq 0. \tag{25}$$

Proof. We shall use the mathematial induction. Let $k \geq 0$ be such that (25) holds. Because $R(A^T)$ is an invariant subspace for the linear application Q (see e.g. [17]), by using (21)-(23) and the induction hypothesis we obtain

$$x^{k+1} = Qx^k + Fb^{k+1} = P_{N(A)}(x^k) + \tilde{Q}x^k + Fb^{k+1}, \tag{26}$$

with $\tilde{Q}x^k + Fb^{k+1} \in R(A^T)$. From (26) it results that $P_{N(A)}(x^{k+1}) = P_{N(A)}(x^k)$ and the proof is complete.
Let now, $\varepsilon^k, e^k, k \geq 0$ be the errors in steps (15) and (17) of the algorithm KECG, defined by (see also Theorem 1 , (12) and (25))

$$\varepsilon^k = y^k - P_{N(A^T)}(b), \quad e^k = x^k - (P_{N(A)}(x^0) + x_{LS}). \tag{27}$$

Lemma 2 *We have the equality*

$$e^k \ = \ \tilde{Q}^k e^0 + \sum_{j=1}^{k-1} \tilde{Q}^{k-j} F \varepsilon^j + F\varepsilon^k, \ \forall k \geq 1. \tag{28}$$

Proof. By using (21) - (25), (27) and the decomposition

$$b = P_{R(A)}(b) + P_{N(A^T)}(b) \tag{29}$$

we successively get, for $k \geq 2$

$$\begin{gathered}
e^k = x^k - (P_{N(A)}(x^0) + x_{LS}) = Qx^{k-1} + Fb^k - (P_{N(A)}(x^0) + x_{LS}) = \\
P_{N(A)}(x^{k-1}) + \tilde{Q}x^{k-1} + Fb^k - (P_{N(A)}(x^0) + x_{LS}) = \tilde{Q}x^{k-1} + Fb^k - x_{LS} = \\
\tilde{Q}x^{k-1} + Fb^k - \left[(I - \tilde{Q}) + \tilde{Q}\right](I - \tilde{Q})^{-1} F P_{R(A)}(b) = \\
\left(\tilde{Q}x^{k-1} - \tilde{Q}GP_{R(A)}(b)\right) + \left(Fb^k - FP_{R(A)}(b)\right) = \tilde{Q}e^{k-1} + F\varepsilon^k.
\end{gathered} \tag{30}$$

Then, a recursive application of (30) gives us (28) and completes the proof.

Theorem 3 *Let us suppose that $A_i \neq 0, i = 1, \ldots, m$. Then, for any $x^0 \in I\!R^n$ the sequence $(x^k)_{k \geq 0}$ generated with the algorithm KECG converges and*

$$\lim_{k \to \infty} x^k = P_{N(A)}(x^0) + x_{LS} \in LSS(A; b). \tag{31}$$

Proof. By taking the euclidean norm in (28), and using (12) and (27) we get

$$\begin{gathered}
\| e^k \| \leq \| \tilde{Q} \|^k \cdot \| e^0 \| + \| F \|_2 \left[\sum_{j=1}^{k-1} \| \tilde{Q} \|^{k-j} \cdot \| \varepsilon^j \| + \| \varepsilon^k \| \right] \leq \\
\| \tilde{Q} \|^k \cdot \| \varepsilon^0 \| + \| F \|_2 \left[\sum_{j=1}^{k-1} \| \tilde{Q} \|^{k-j} \cdot \gamma \cdot \delta^j + \gamma \cdot \delta^k \right],
\end{gathered} \tag{32}$$

where by $\| F \|_2$ we denoted the spectral norm of the matrix F. Now, if we denote by α the number

$$\alpha = \max\{\| \tilde{Q} \|, \ \delta\} \in [0, 1), \tag{33}$$

from (32) we get

$$\| e^k \| \leq \ \alpha^k \cdot \| e^0 \| + \| F \|_2 \, \dot{\gamma} \cdot k \cdot \alpha^k, \ \forall k \geq 1 \tag{34}$$

and from (33) $\lim_{k \to \infty} e^k = 0$, which is exactly (31) and the proof is complete.
If we replace the projections $f_i(\beta; x)$ in (5) with

$$f_i(\omega; \beta; x) = (1 - \omega)x + \omega f_i(\beta; x) = x - \omega \frac{\langle x, A_i \rangle - \beta_i}{\| A_i \|^2} A_i, \tag{35}$$

we can define the KECG algorithm with Relaxation Parameter as follows.
Algorithm KECGRP. Let $x^0 \in I\!R^n, y^0 = b$; for $k = 0, 1, \ldots$ do

$$\begin{aligned} y^{k+1} &= GCG(A^T; y^k) && (36) \\ b^{k+1} &= b - y^{k+1} && (37) \\ x^{k+1} &= (f_1 \circ \ldots \circ f_m)(\omega; b^{k+1}; x^k) && (38) \end{aligned}$$

Then, according to the above Theorem 3, the considerations from section 4 in [14] directly apply and give us the following convergence result.

Theorem 4 *If A is as in the hypothesis of Theorem 3, for any $x^0 \in I\!R^n$ and $\omega \in (0, 2)$, the sequence $(x^k)_{k \geq 0}$ generated with the algorithm KECGRP converges and*

$$\lim_{k \to \infty} x^k = P_{N(A)}(x^0) + x_{LS} \in LSS(A; b). \qquad (39)$$

Remark 3 *In the paper [5], T. Elfving proposed the following generalization of the algorithm KE (2) − (4).*
Algorithm TELF. *Let $x^0 \in I\!R^n, y^0 = b - Ax^0$; for $k = 0, 1, \ldots$ do*

$$\begin{aligned} y^{k+1} &= (I - AR)y^k && (40) \\ x^{k+1} &= \bar{Q}x^k + \bar{R}(b - y^{k+1}), && (41) \end{aligned}$$

where the matrices $R, \bar{Q}, \bar{R}$ are constructed such that several properties hold for the two steps (40) − (41) of the algorithm. We want to observe that in our algorithm KECG, the step (15) cannot be of the form (40) (because of the special structure of the GCG algorithm). Thus, algorithm KECG is not particular case of TELF. They are both methods for solving inconsistent problems like (1), but they can not be obtained one from the other because one iteration of GCG algorithm cannot be written in the form (40) and the methods mentioned in [5] do not always satisfy step error reduction formulas like (12).

4 Numerical experiments

We considered the original image from Figure 2, $x^{ex} \in I\!R^{1024 \times 1024}$ and the consistent system

$$Ax = b^0, \qquad (42)$$

where the matrix A was constructed following the fan-beam scanning procedure indicated in Figure 3, with 1140 sources S_i and 2048 detectors D_j for each position of the source S_i, and $b^0 = Ax^{ex}$. Because A is a (sparse) matrix of dimensions m=1140*2048=2.334.720 and n=1024*1024=1.048.576, it couldn't be anymore stored in the computer memory (Intel Pentium 4, 3.00GHz, 1GB RAM). We then applied to the corresponding system (42) the KE and GCG algorithms as follows.

- GCG: we (re)generated the rows of A two times in each iteration and performed the computations as in Example 1 before;
- KE: for step (4) we (re)generated the rows of A one time in each iteration, whereas for step (2) we constructed and stored on the disk (in a pre-processing phase) the transpose A^T and then, in each iteration we successively read from it the columns;

The reconstructed images, obtained after 5 iterations of each algorithm are presented in Figures 4 and 5, in which we also indicate the computational time. We can observe in Figure 5 the blurring effect created by the GCG algorithm (as mentioned in [8]; see also [7]).
We then simulated a perturbation of the right hand side of (42) according to the formula

$$b = b^0 + pert, \quad pert = 10\% \cdot \parallel b^0 \parallel \cdot rand, \qquad (43)$$

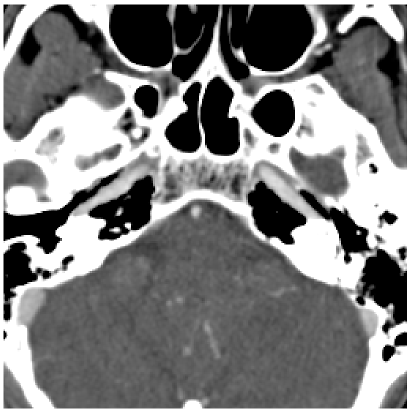

Figure 2: Original image

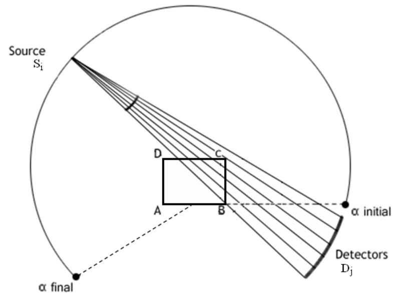

Figure 3: Fan beam scanning

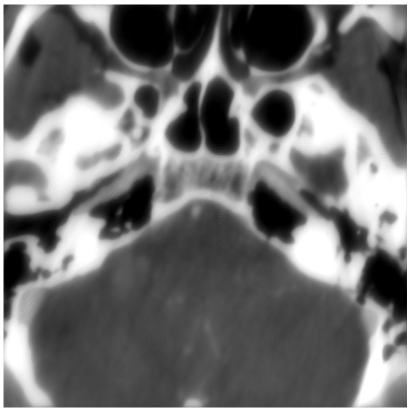

Figure 4: CG - consistent Time: 70 min.

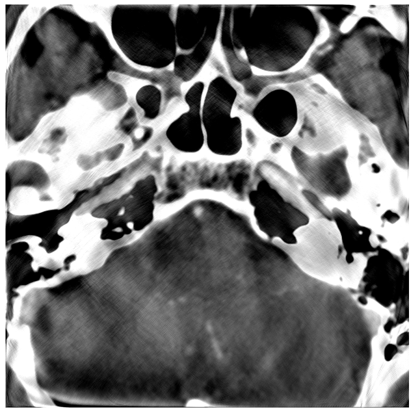

Figure 5: KE - consistent Time: 900 min.

with *rand* a randomly generated vector with unitary norm and we considered the Tikhonov-like regularization problem associated to (42)

$$\| Ax - b \|^2 + \delta^2 \| Lx \|^2 = \min! \Leftrightarrow \left\| \left[\begin{array}{c} A \\ \delta L \end{array} \right] x - \left[\begin{array}{c} b \\ 0 \end{array} \right] \right\|^2 = \min!, \tag{44}$$

where L is an $n \times n$ matrix constructed according to the following procedure (see e.g. [1] and references therein): for each $i \in 1, \ldots, n$, let H_i be the set of horizontally neighbour pixels of i, V_i the set of vertically neighbour pixels and D_i the set of diagonally neighbour pixels of i. Then, for each $j \in 1, \ldots, n$ we define $(L)_{ij}$ as

$$(L)_{ij} = \begin{cases} (L)_{ij} = w_h, & \text{if } j \in H_i \\ (L)_{ij} = w_v, & \text{if } j \in V_i \\ (L)_{ij} = w_d, & \text{if } j \in D_i \\ \sum_{k=1}^{n} |(L)_{ik}|, & \text{if } j = i \text{ and } k \neq i \\ 0, & \text{otherwise} \end{cases} \tag{45}$$

In our experiments we used $w_h = -1, w_v = -1, w_d = -1/\sqrt{2}$ and the same procedure as we described before for KE and GCG with respect to the usage of the regularized matrix from (44) in each iteration of KE and KECG. In Figures 6 and 7 we present the reconstructed images after 10 iterations of the algorithms KE and KECG applied to the regularized problem (44), with L defined as in (45) and $\delta^2 = 20$, by also indicated the computational time. According to these figures, we can observe that the image reconstructed with the new KECG algorithm is very close to the KE one, although - as we expected - this is better. On the other hand, the computational time used for KE is much bigger than for KECG.

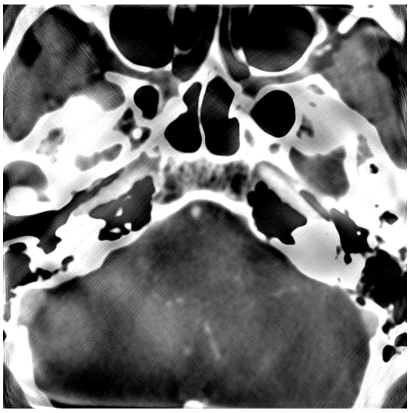

Figure 6: KE - regularized Time: 1010 min.

Figure 7: KECG - regularized Time: 105 min.

Final remarks and future work. In this paper we constructed and theoretically analysed a hybrid algorithm for computation of the minimal norm solutions of inconsistent and rank-deficient least squares problems. It combines in one iteration a CG-like step together with a Kaczmarz-like one. Numerical experiments performed on a real 2D image arising from medical applications, show that our algorithm is very close to the Kaczmarz-Extended one. On the other hand, there is a computational time advantage of the new algorithm against KE, which makes it possible to be used for (very) large scale reconstruction problems. The main objective for further developments would be the optimization of the implementation of the new algorithm, such that it becomes efficient in real applications.

Acknowledgements. The original image from figure 2 has been provided by prof. dr. Joachim Hornegger and dr. Marcus Prümmer from University of Erlangen-Nürnberg, Germany. The software for the numerical experiments has been designed by my master student Tiberius Duluman, from Ovidius University of Constanta, Romania. I express my gratitude to all of them.

References

[1] Bautu A., Bautu E., Popa C., *Tikhonov regularization in image reconstruction with Kaczmarz Extended algorithm*, Proceedings of ASIM Conference, September 12-15 2005, Erlangen, Germany, SCS Publishing House e.V., 650-655.

[2] Björck, A., *Numerical methods for least squares problems*, SIAM Philadelphia, 1996.

[3] Censor Y., Stavros A. Z., *Parallel optimization: theory, algorithms and applications*, "Numer. Math. and Sci. Comp." Series, Oxford Univ. Press, New York, 1997.

[4] Censor Y., *Row-action methods for huge and sparse systems and their applications*, SIAM Review, **23(4)** (1981), pp. 444-466.

[5] Elfving T., *A stationary iterative pseudoinverse algorithm*, *B I T*, **38(2)**(1998), 275–282.

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

[7] Herman, G. T., *Image reconstruction from projections. The fundamentals of computerized tomography*, Academic Press, New York, 1980.

[8] Herman, G. T., *Personal communication*, December 2005.

[9] Kaczmarz S., *Angenaherte Auflosung von Systemen linearer Gleichungen*, Bull. Acad. Polonaise Sci. et Lettres **A** (1937), 355–357.

[10] Kammerer W. J., Nashed M. Z., *On the convergence of the conjugate gradient method for singular linear operator equations*, SIAM J. Numer. Anal., **9**(1972), 165–181.

[11] Koltracht I., Lancaster P. and Smith D., *The structure of some matrices arising in tomography*, Linear Algebra and its Applic., **130**(1990), 193–218.

[12] Natterer F., *The Mathematics of Computerized Tomography*, John Wiley and Sons, New York, 1986.

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

[14] Popa C., *Extensions of block-projection methods with relaxation parameters to inconsistent and rank-defficient least-squares problems*, *B I T*, **38(1)**(1998), 151–176.

[15] Popa C. and Zdunek R., *Kaczmarz extended algorithm for tomographic image reconstruction from limited-data*, Math. and Computers in Simulation, **65**(2004), 579–598.

[16] Prümmer, M., Köstler, H. et al., *A full multigrid technique to accelerate an ART scheme for tomographic image reconstruction*, Proceedings of ASIM Conference, September 12-15 2005, Erlangen, Germany, SCS Publishing House e.V., 632-637.

[17] Tanabe K., *Projection Method for Solving a Singular System of Linear Equations and its Applications*, Numer. Math., **17**(1971), 203–214.

[18] Tanabe K., *Conjugate -gradient method for computing the Moore - Penrose inverse and rank of a matrix*, Journal of Optim. Theory and Appl., **22(1)**(1977), 1–23.