

# Kaczmarz Extended versus augmented system solution in image reconstruction

Aurelian Nicola, Constantin Popa

“Ovidius” University of Constanta, Romania,  
e-mail: {anicola, cpopa}@univ-ovidius.ro

**Abstract.** In this paper we make a comparative analysis of two projection based iterative algorithms for systems of linear equations arising from image reconstruction in computerized tomography. The first one - Kaczmarz’s iterations - is used for solving the (consistent) augmented system, whereas the second - Kaczmarz extended algorithm - is used for solving the original (inconsistent) system. We obtain bounds for the generalized spectral condition numbers for both augmented and original system, which give us information about the theoretical behaviour of the two iterative methods. Numerical experiments are made on a phantom used in the image reconstruction literature.

## 1 Introduction

Image reconstruction from projections in computerized tomography gives rise to general least squares problem (LSP, for short): find  $x^* \in \mathbb{R}^n$  such that

$$\| Ax^* - b \| = \min\{\| Az - b \|, z \in \mathbb{R}^n\}, \quad (1)$$

where  $\langle \cdot, \cdot \rangle, \| \cdot \|$  will denote the Euclidean scalar product and norm. We shall denote by  $LSS(A; b)$  the set of its all (least squares) solutions and by  $x_{LS}$  its (unique) minimal norm one.  $A^T, \mathcal{R}(\mathcal{A}), \mathcal{N}(\mathcal{A}), \sigma(A), \rho(A)$  will denote the transpose, range, null space, spectrum and spectral radius of  $A$ , whereas  $P_S$  will be the orthogonal projection onto a vector subspace  $S \in \mathbb{R}^q$ , for some  $q \geq 1$ . Moreover, if  $r = rank(A)$  and

$$U^T AV = diag(\sigma_1, \dots, \sigma_r, 0, \dots, 0), \quad \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0 \quad (2)$$

is a singular value decomposition (SVD) of  $A$  we define its *generalized spectral condition number* by

$$gk_2(A) = \frac{\sigma_1}{\sigma_r}. \quad (3)$$

**Remark 1** If  $B$  is a symmetric  $n \times n$  matrix with  $r = \text{rank}(B)$ , it exists an orthogonal one  $U$  such that

$$U^T B U = \text{diag}(\lambda_1, \dots, \lambda_n), \quad (4)$$

where  $\lambda_i \in \mathbb{R}$  are the eigenvalues of  $B$ . If in addition  $\lambda_i \geq 0, \forall i = 1, \dots, n$  (i.e.  $B$  is positive semidefinite) then we can construct  $U$  such that

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0 = \lambda_{r+1} = \dots = \lambda_n,$$

which tells us that the decomposition (4) is an a SVD of  $B$  and

$$gk_2(B) = \frac{\lambda_1}{\lambda_r}. \quad (5)$$

From all these considerations and (2) we get for a general matrix the equality

$$gk_2(A) = \sqrt{gk_2(A^T A)} = \sqrt{\frac{\lambda_{\max}(A^T A)}{\lambda_{\min}^*(A^T A)}}, \quad (6)$$

where  $\lambda_{\max}(A^T A), \lambda_{\min}^*(A^T A)$  are the maximal, respectively minimal-nonzero eigenvalues of the symmetric and positive semidefinite matrix  $A^T A$ .

The following results are known.

**Proposition 1** We have the following two equivalent formulations for (1):

$$x^* \in LSS(A; b) \iff Ax^* = P_{\mathcal{R}(A)}(b). \quad (7)$$

and the normal equation

$$x^* \in LSS(A; b) \iff A^T A x^* = A^T b. \quad (8)$$

**Remark 2** The normal equation (8) is not only a theoretical characterization of  $LSS(A; b)$ , but because it is square, positive semidefinite and always consistent, several direct and iterative solvers have been designed for its solution (see e.g. [1, 2]). Although the most of these methods do not require the computation of  $A^T A$ , a big inconvenient still remains: the generalized spectral condition number for the matrix involved in the normal equation satisfies (see (6))  $gk_2(A^T A) = (gk_2(A))^2$ , which will determine a poor behaviour of the corresponding methods when applied to (8).

As an alternative with respect to the inconvenients mentioned in the above remark, we may consider the associated  $(m+n) \times (m+n)$  augmented system (see [1])

$$\begin{bmatrix} I & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} r \\ x \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}. \quad (9)$$

**Proposition 2** (i) *The normal equation (8) and the augmented system (9) are equivalent in the following sense: if  $x \in \mathbb{R}^n$  satisfies (8) and  $r = b - Ax$ , then  $(r, x)^T \in \mathbb{R}^{m+n}$  satisfies (9) and conversely, if  $(r, x)^T \in \mathbb{R}^{m+n}$  is a solution for (9) then  $x$  satisfies (8). In particular, the augmented system is consistent.*

(ii) *Let  $x_{LS}$  be the minimal norm solution of (1) and*

$$r_{LS} = b - Ax_{LS} = P_{\mathcal{N}(A^T)}(b). \quad (10)$$

*Then, the minimal norm solution of the augmented system  $z_{LS} \in \mathbb{R}^{m+n}$  is given by*

$$z_{LS} = (r_{LS}, x_{LS})^T. \quad (11)$$

**Proof.** (i) If  $x \in \mathbb{R}^n$  is a solution for (8) and  $r = b - Ax$  then

$$\begin{bmatrix} I & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} b - Ax \\ x \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix},$$

i.e.  $(r, x)^T$  is a solution for (9). Conversely, if  $(r, x)^T \in \mathbb{R}^{m+n}$  is a solution for (9) then

$$\begin{cases} r + Ax = b \\ A^T r = 0 \end{cases} \Rightarrow \begin{cases} A^T r + A^T Ax = A^T b \\ A^T r = 0 \end{cases},$$

i.e.  $x$  is a solution of (8).

A let (ii) s in (i) we obtain that  $z_{LS}$  from (11) is among the solutions of the augmented system (9). Then, let  $z = (r, x)^T \in \mathbb{R}^{m+n}$  be an arbitrary solution of (9). As before it results that

$$x \in LSS(A; b) \Leftrightarrow Ax = P_{\mathcal{R}(A)}(b), \text{ thus } r = b - Ax = P_{\mathcal{N}(A^T)}(b), \quad (12)$$

where we have also used the orthogonal decomposition  $b = P_{\mathcal{R}(A)}(b) + P_{\mathcal{N}(A^T)}(b)$ . Then, we successively get, by also using (10) and (12)

$$\begin{aligned} \|z\|^2 &= \|(r, x)^T\|^2 = \|P_{\mathcal{N}(A^T)}(b)\|^2 + \|x\|^2 \geq \\ &\|P_{\mathcal{N}(A^T)}(b)\|^2 + \|x_{LS}\|^2 = \|(r_{LS}, x_{LS})^T\|^2 = \|z_{LS}\|^2 \end{aligned}$$

and the proof is complete.

**Proposition 3** *If  $M$  is the  $(m+n) \times (m+n)$  symmetric matrix of the augmented system (9), i.e.*

$$M = \begin{bmatrix} I & A \\ A^T & 0 \end{bmatrix} \quad (13)$$

then

$$gk_2(M) = \mathcal{O}((gk_2(A))^2). \quad (14)$$

**Proof.** The first part of the proof is adapted from [1]. Let  $\lambda \in \sigma(M)$  be an eigenvalue and  $0 \neq z = (s, x)^T \in \mathbb{R}^{m+n}$  a corresponding eigenvector, i.e.  $Mz = \lambda z$  thus, by using (13)

$$s + Ax = \lambda s, \quad A^T s = \lambda x. \quad (15)$$

If we multiply from the left with  $A^T$  the first equation in (15) and we then replace  $A^T s$  by  $\lambda x$  from the second one, we get the equality

$$A^T Ax = (\lambda^2 - \lambda)x, \quad (16)$$

which tells us that, if  $x \neq 0$  then

$$\lambda^2 - \lambda \in \sigma(A^T A). \quad (17)$$

If  $x = 0$ , then  $s \neq 0$  and from the first equality in (15) we obtain

$$s = \lambda s \Leftrightarrow \lambda = 1. \quad (18)$$

But, according to (2) we have

$$\sigma(A^T A) = \{\sigma_1^2, \dots, \sigma_r^2, 0\}. \quad (19)$$

Then, by using (17) - (19) we obtain

$$\sigma(M) = \left\{ \frac{1}{2} \pm \sqrt{\frac{1}{4} + \sigma_i^2}, i = 1, \dots, r \right\} \cup \{1, 0\} \quad (20)$$

where  $\sigma_1 \geq \dots \geq \sigma_r > 0$  are the singular values of  $A$  and the eigenvalues 1 and 0 have the multiplicity  $m - r$  and  $n - r$ , respectively.

In the second part of the proof we will evaluate  $gk_2(M)$  according to Remark 1 formula (6). From (20) and because  $M^T M = M^2$  we get

$$\frac{1}{2} + \sqrt{\frac{1}{4} + \sigma_i^2} > 1, \quad \forall i = 1, \dots, r \quad (21)$$

and

$$\left| \frac{1}{2} - \sqrt{\frac{1}{4} + \sigma_i^2} \right| = \frac{\sigma_i^2}{\frac{1}{2} + \sqrt{\frac{1}{4} + \sigma_i^2}} = \mathcal{O}(\sigma_i^2), \quad (22)$$

thus, it will result from (2)

$$\lambda_{min}^*(M^T M) = \left( \frac{1}{2} - \sqrt{\frac{1}{4} + \sigma_r^2} \right)^2, \lambda_{max}(M^T M) = \left( \frac{1}{2} + \sqrt{\frac{1}{4} + \sigma_1^2} \right)^2. \quad (23)$$

Then, according to (6) we obtain

$$gk_2(M) = \sqrt{\frac{\lambda_{max}(M^T M)}{\lambda_{min}^*(M^T M)}} = \frac{\frac{1}{2} + \sqrt{\frac{1}{4} + \sigma_1^2}}{\sqrt{\frac{1}{4} + \sigma_r^2} - \frac{1}{2}} = \frac{\left(\frac{1}{2} + \sqrt{\frac{1}{4} + \sigma_1^2}\right)\left(\frac{1}{2} + \sqrt{\frac{1}{4} + \sigma_r^2}\right)}{\sigma_r^2}. \quad (24)$$

But, for any  $x > 0$  we have

$$1 + x > \frac{1}{2} + \sqrt{\frac{1}{4} + x^2} > \frac{1}{2} + x \quad (25)$$

which combined with (24) gives us

$$\frac{(1 + \sigma_1)^2}{\sigma_r^2} > \frac{(1 + \sigma_1)(1 + \sigma_r)}{\sigma_r^2} > gk_2(M) > \frac{\left(\frac{1}{2} + \sigma_1\right)\left(\frac{1}{2} + \sigma_r\right)}{\sigma_r^2} > \frac{1}{4} \frac{1}{\sigma_r^2}. \quad (26)$$

Because of (3) and the fact that we may suppose without restricting the generality of the problem that  $\sigma_1 = \mathcal{O}(1)$ , from (26) we get (14) and the proof is complete.

## 2 Kaczmarz and Kaczmarz Extended algorithms

First proposed by its author in [5], Kaczmarz's projection method has been developed by many others in various directions (see [1, 2, 6, 7, 9] and references therein). If we consider the applications  $f_i(b; \cdot), F(b; \cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , defined by

$$f_i(\omega; b; x) = x - \omega \frac{\langle x, A_i \rangle - b_i}{\|A_i\|^2} A_i, \quad F(\omega; b; x) = (f_1 \circ \dots \circ f_m)(\omega; b; x), \quad (27)$$

where, for the moment  $\omega \neq 0$  and  $b \in \mathbb{R}^m$  are considered as fixed parameters in  $f_i(\omega; b; \cdot)$ , then Kaczmarz's iteration with relaxation parameter ( $\omega - K$  for short) for the problem (1) can be written as follows.

**Algorithm  $\omega - K$ .** *Initialization:*  $x_0 \in \mathbb{R}^n$

*Iterative step:*

$$x^{k+1} = F(\omega; b; x^k), \quad k \geq 0. \quad (28)$$

**Proposition 4** (see [2, 6]) *If the problem (1) is consistent, for any  $\omega \in (0, 2)$  and  $x^0 \in \mathbb{R}^n$  the sequence  $(x^k)_{k \geq 0}$  generated by (28) converges and*

$$\lim_{k \rightarrow \infty} x^k = P_{N(A)}(x^0) + x_{LS} \in S(A; b). \quad (29)$$

**Remark 3** *Let*

$$M \begin{bmatrix} r \\ x \end{bmatrix} = \hat{b}, \quad \hat{b} = [b, 0]^T \quad (30)$$

*be the augmented system (9) (with  $M$  from (13)). The above Proposition 4 tell us that the  $\omega$ -Kaczmarz algorithm (28) applied to the augmented system (30), with the initial approximation  $(r^0, x^0)^T = (0, 0)^T$  converges to the minimal norm solution ( $z_{LS}$  from (11)) of the augmented system (30).*

Unfortunately, in the inconsistent case for (1) (which appears in practical applications) the above Kaczmarz's sequence  $(x^k)_{k \geq 0}$  from (28) still converges, but its limit is at a certain distance from the set  $LSS(A; b)$  (see e.g. [8]). For overcoming this difficulty, one of the authors proposed in [7] the following extension of the algorithm  $\omega - K$ .

**Algorithm Kaczmarz Extended with Relaxation Parameters (KERP).**

*Initialization:*  $\alpha, \omega \in (0, 2)$ ;  $x_0 \in \mathbb{R}^n$ ,  $y^0 = b$ ;

*Iterative step:*

$$y^{k+1} = \Phi(\alpha, y^k), \quad (31)$$

$$b^{k+1} = b - y^{k+1}, \quad (32)$$

$$x^{k+1} = F(\omega; b^{k+1}; x^k). \quad (33)$$

with

$$\Phi(\alpha, y) = (\varphi_1 \circ \dots \circ \varphi_n)(\alpha; y), \quad \varphi_j(\alpha; y) = (1 - \alpha)y + \alpha\varphi_j(y) \quad (34)$$

and

$$\varphi_j(y) = y - \frac{\langle y, A^j \rangle}{\|A^j\|^2} A^j, \quad j = 1, \dots, n. \quad (35)$$

**Proposition 5** For any problem of the form (1), any  $\alpha, \omega \in (0, 2)$  and  $x^0 \in \mathbb{R}^n$  the sequence  $(x^k)_{k \geq 0}$  generated by the algorithm KERP (31)-(33) converges and

$$\lim_{k \rightarrow \infty} x^k = P_{N(A)}(x^0) + x_{LS} \in LSS(A; b). \quad (36)$$

As proved in [6], the sequence  $(x^k)_{k \geq 0}$  obtained by applying  $\omega - K$  algorithm to (28) can be constructed as  $x^k = A^T y^k$ , where the sequence  $(y^k)_{k \geq 0}$  is constructed by applying the SOR iterative method (see [10]) to the system

$$AA^T y = b. \quad (37)$$

Then, we expect for  $\omega - K$  algorithm a behaviour as for the SOR iteration applied to (37), i.e. an asymptotic convergence rate depending on

$$gk_2(AA^T) = gk_2(A^T A) = (gk_2(A))^2. \quad (38)$$

Thus, if  $\omega - K$  will be applied to the augmented system (9), we expect an asymptotic convergence rate depending on (see (38) and (14))

$$gk_2(MM^T) = gk_2(M^T M) = (gk_2(M))^2 = \mathcal{O}((gk_2(A))^4). \quad (39)$$

On the other hand, because the step (31) of the algorithm KERP is an  $\omega - K$  like iteration applied to the system  $A^T y = 0$  and the step (33) is an  $\omega - K$  iteration applied to the system  $Ax = b^{k+1}$ , according to (38) we expect for the whole KERP method (31)-(33) a convergence rate depending on the condition numbers in (38), i.e. finally of order  $\mathcal{O}((gk_2(A))^2)$ . Then, relations (38) and (39), together with the above comments tell us that we expect that the KERP algorithm applied to the initial problem (1) will be faster and will produce better reconstruction than the  $\omega - K$  algorithm applied to the augmented system (9).

### 3 Experiments

We have used in our experiments the *mitochondrian phantom* from the paper [3] (see also [4] for a complete description of a phantom). It contains the exact picture  $x^{ex} \in \mathbb{R}^{3969}$ , i.e. with a  $63 \times 63$  pixels resolution from Figure 1 left, a scanning matrix  $A : 1378 \times 3969$  and a measurements right hand side  $b \in \mathbb{R}^{1378}$ . The associated reconstruction problem is an inconsistent formulation

as (1). We applied 100 iterations with the algorithms from Section 2 (with  $\alpha = \omega = 1$ ), as it is described there. The reconstructions are presented in Figure 1. Figure 2 shows the wellknown four error measures used in image reconstruction (see e.g. [4]): standard deviation, distance, relative error and normal equation residual, defined below.

- $x^{ex} = (x_1^{ex}, \dots, x_{3969}^{ex})^T$  - the mitochondrial phantom
- $x^k = (x_1^k, \dots, x_{3969}^k)^T$  - the current approximation
- $\bar{x}^{ex} = \frac{1}{3969} \sum_{i=1}^{3969} x_i^{ex}$ ;  $\bar{x}^k = \frac{1}{3969} \sum_{i=1}^{3969} x_i^k$  - mean values of the exact image and current approximation, respectively
- **Standard deviation** =  $\frac{1}{\sqrt{3969}} \sqrt{\sum_{i=1}^{3969} (x_i^k - \bar{x}^k)^2}$
- **Distance** =  $\sqrt{\frac{\sum_{i=1}^{3969} (x_i^{ex} - x_i^k)^2}{\sum_{i=1}^{3969} (x_i^{ex} - \bar{x}^{ex})^2}}$
- **Relative error** =  $\frac{\sum_{i=1}^{3969} |x_i^{ex} - x_i^k|}{\sum_{i=1}^{3969} x_i^{ex}}$
- **Normal equation residual** =  $\frac{1}{63} \| A^T(Ax^k - b) \|$

The results presented in Figure 2 show the better behaviour of KERP algorithm.

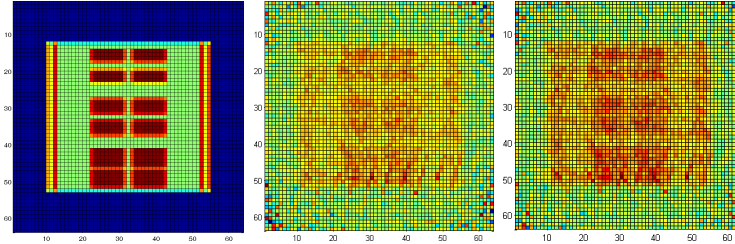


Figure 1: Exact image,  $\omega$ -Kaczmarz for (30) , KERP for (1)



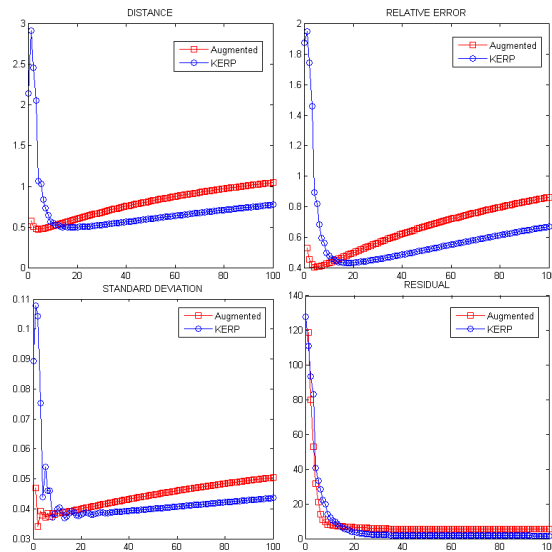


Figure 2: Top-left: distance, top-right: relative error, bottom-left: standard deviation, bottom-right: normal equation residual

## References

- [1] Bjork A., *Numerical methods for least squares problems*, SIAM Philadelphia, 1996.
- [2] Censor Y., Stavros A. Z., *Parallel optimization: theory, algorithms and applications*, "Numer. Math. and Sci. Comp." Series, Oxford Univ. Press, New York, 1997.
- [3] Censor, Y., Elfving T., Herman G. T., Nikazad T., *On diagonally-relaxed orthogonal projection methods*, SIAM J. Sci. Comput., **30**(2007/08), 473–504.
- [4] Herman, G. T., *Image reconstruction from projections. The fundamentals of computerized tomography*, Academic Press, New York, 1980.
- [5] Kaczmarz S., *Angenaherte Auflosung von Systemen linearer Gleichungen*, Bull. Acad. Polonaise Sci. et Lettres **A** (1937), 355–357.
- [6] Natterer F., *The Mathematics of Computerized Tomography*, John Wiley and Sons, New York, 1986.

- [7] Popa C., *Extensions of block-projections methods with relaxation parameters to inconsistent and rank-deficient least-squares problems*; *B I T*, **38(1)**(1998), 151–176.
- [8] Popa C. and Zdunek R., *Kaczmarz extended algorithm for tomographic image reconstruction from limited-data*, *Math. and Computers in Simulation*, **65**(2004), 579–598.
- [9] Tanabe K., *Projection Method for Solving a Singular System of Linear Equations and its Applications*, *Numer. Math.*, **17**(1971), 203–214.
- [10] Young D. M., *Iterative solution of large linear systems*, Academic Press, New York, 1971.