Tikhonov Regularization in Image Reconstruction with Kaczmarz Extended Algorithm[1]

Andrei Băutu
abautu@anmb.ro
"Mircea cel Bătrân" Naval Academy
1^{st} Fulgerului St, Constantza, 900218
Romania

Elena Băutu, Constantin Popa
{erogojina, cpopa}@univ-ovidius.ro
"Ovidius" University
124^{th} Mamaia Blvd., Constantza, 900527
Romania

Abstract

In a previous paper we proposed a simple and natural extension of Kaczmarz's projection algorithm (KE, for short) to inconsistent least-squares problems arising in ART image reconstruction in computerized tomography. In the present one we describe two versions of this extension for a Tikhonov regularization of the original inconsistent least-squares problem. The first version deals directly with an (augmented) equivalent formulation of the Tikhonov regularization problem, whereas the second one uses the gradient of the Tikhonov regularized functional. For both new versions of the KE algorithm we present some theoretical considerations together with numerical experiments and comparisons with the initial KE method.

1 Introduction

Many problems in the field of tomographic image reconstruction are modeled by the linear least-squares problem: find $x \in I\!R^n$ such that

$$\|Ax - \tilde{b}\| = \min!, \tag{1}$$

where A is an $m \times n$ real matrix and $\tilde{b} \in I\!R^m$ a given vector ($\| \cdot \|$ and $\langle \cdot, \cdot \rangle$ will denote the Euclidean norm and scalar product on some space $I\!R^q$). Although from a theoretical view point the problem (1) is consistent, i.e $\tilde{b} \in R(A)$, in real world applications, usually due to measurements errors, the right hand side of (1) is perturbed as

$$b = \tilde{b} + \delta b, \ \ \delta b = \delta b_A + \delta b_A^* \in R(A) \oplus N(A^t) \tag{2}$$

and the problem becomes inconsistent ($R(A), N(A), A^t$ will denote the range, null space and transpose of A). In this case, the classical Kaczmarz projection algorithm (see [6]) can no longer be used, thus the extended version KE from [5] has to be applied. If $a_i \in$

[1] Paper supported by the PNCDI INFOSOC Grant 131/2004

$I\!R^n, \alpha_j \in I\!R^m$, denote the i-th row and j-th column of A, respectively (which we suppose to be nonzero vectors) the KE algoritm for the problem (1), with $\tilde{b}$ replaced by b, can be written as follows.
Algorithm KE. Let $x^0 \in I\!R^n, y^0 = b$; for $k = 0, 1 \ldots$ do

$$y^{k+1} = \Phi(\alpha; y^k); \;\; b^{k+1} = b - y^{k+1}; \;\; x^{k+1} = F(\omega; b^{k+1}; x^k). \tag{3}$$

Here α, ω are relaxation parameters and the applications involved in (3) are defined by (see [5] for details)

$$f_i(\omega; b; x) = (1-\omega)x + \omega f_i(b; x), \;\; F(\omega; b; x) = (f_1 \circ \ldots \circ f_m)(\omega; b; x), \tag{4}$$

$$\varphi_j(\alpha; y) = (1-\alpha)y + \alpha\varphi_j(y), \;\; \Phi(\alpha; y) = (\varphi_1 \circ \ldots \circ \varphi_n)(\alpha; y), \tag{5}$$

$$f_i(b; x) = x - \frac{\langle x, a_i\rangle - b_i}{\|a_i\|^2} a_i, \;\; \varphi_j(y) = y - \frac{\langle y, \alpha_j\rangle}{\|\alpha_j\|^2}\alpha_j. \tag{6}$$

According to the perturbation in (2), the above KE algorithm has the following property.

Theorem 1 *If $\tilde{b} \in R(A)$, then $b^k \in R(A), \forall k \geq 0$, where $(b^k)_{k\geq 0}$ is the sequence generated during the KE algorithm (3).*

Proof. If $z \in I\!R^m$ is a vector from $N(A^t)$, then $\langle z, \alpha_j\rangle = 0, \forall j = 1, \ldots, n$ so (see (5)) $\Phi(\alpha; z) = z$. Thus, in the first KE iteration (see also (2)) we obtain $\Phi(\alpha; y^0) = \Phi(\alpha; b) = \Phi(\alpha; b + \delta b_A) + \delta b_A^*$, i.e. $b^1 = b - y^0 = b - \Phi(\alpha; b) = \tilde{b} + \delta b_A \in R(A)$ and a recursive argument completes the proof.
But, (according to the above result) although the $N(A^t)$ component of the perturbation vector δb from (2) is completely eliminated during the KE algorithm, the other one, $\delta b_A \in R(A)$ can still play an unpleasant role (see e.g. the examples in [6]). In order to eliminate also this bad influence, we have to consider a Tikhonov type regularization (for the perturbed problem $\|Ax - b\| = \min!$) of the form (see e.g. [2])

$$\|Ax - b\|^2 + \gamma^2\langle Rx, x\rangle = \min! \tag{7}$$

where $\gamma \in (0, \infty)$ and R is an $n \times n$ symmetric and positive semidefinite matrix constructed with the local information involving neighbours pixels in the domain (picture) discretization (this construction will be described in Section 3 of the paper). The problem (7) is still inconsistent, thus we have to apply for it a corresponding algorithm. In this respect, two versions of the above KE algorithm will be described in the next section.

2 The regularized extended Kaczmarz algorithms

The first version of the KE algorithm for the regularized problem (7), which will be denoted by RKE-1, is based on the simple observation that, if the regularization matrix R is in addition positive definite it will have a Cholesky decomposition of the form $R = LL^t$ and the regularized formulation (7) will be equivalent with

$$\|\hat{A}x - \hat{b}\| = \min!, \tag{8}$$

where

$$\hat{A} = \left\| \begin{array}{c} A \\ \gamma L^t \end{array} \right\|, \quad \hat{b} = \left\| \begin{array}{c} b \\ 0 \end{array} \right\|, \tag{9}$$

with $\hat{A} : (m+n) \times n$ and $\hat{b} \in I\!R^{m+n}$. Then, the RKE-1 algorithm will be exactly KE written for the problem (8)-(9).

Algorithm RKE-1. Let $x^0 \in I\!R^n, \hat{y}^0 = \hat{b}$; for $k = 0, 1 \dots$ do

$$\hat{y}^{k+1} = \hat{\Phi}(\alpha; \hat{y}^k); \;\; \hat{b}^{k+1} = \hat{b} - \hat{y}^{k+1}; \;\; x^{k+1} = \hat{F}(\omega; \hat{b}^{k+1}; x^k), \tag{10}$$

where the applications $\hat{\Phi}, \hat{F}$ are defined as in (4)-(6), but with respect to $\hat{A}, \hat{b}$ from (9). From [5] we can derive the following convergence result for the algorithm RKE-1.

Theorem 2 *For any $x^0 \in I\!R^n$ and any $\omega, \alpha \in (0, 2)$, the sequence $(x^k)_{k \geq 0}$ generated by the algorithm* RKE-1 *converges always to a least-squares solution of the regularized problem $(8) - (9)$. Moreover, for $x^0 = 0$ the limit is exactly $x_{LS}(\gamma)$, i.e the minimal norm solution of $(8) - (9)$.*

The second extension, was inspired by the construction from [4] (see also [6]). In this sense we consider, instead of (7) the following regularized formulation

$$\Psi(x) = \frac{1}{2}\|Ax - b\|_W^2 + \frac{1}{2}\gamma^2 \langle Rx, x \rangle = \min!, \tag{11}$$

where $W = \operatorname{diag}\left(\frac{1}{\|a_1\|^2}, \dots, \frac{1}{\|a_m\|^2}\right)$. Then, all the minimizers of the functional Ψ, i.e. solutions of the regularized problem (11) satisfy the normal equation $\nabla\Psi(x) = 0$, where

$$\nabla\Psi(x) = A^t(Ax - b) + \gamma^2 Rx. \tag{12}$$

This allows as to a simultaneous (Landweber-like) iterative algorithm of the form

$$x^{k+1} = x^k - \lambda_k \left(A^t W(Ax^k - b) + \gamma^2 Rx^k \right), \tag{13}$$

where λ_k is a relaxation parameter, generally depending on the iteration index k . Then, by using the ideas and procedure from the papers [4] and [6] and taking $\lambda_k = 1, \forall k \geq 0$, we proposed the following successive version of (13) (corresponding to a Kaczmarz-like iteration for (11); see also (4))

$$x^{k+1} = F(\omega; b; x^k) - \gamma^2 Rx^k. \tag{14}$$

This allows us to define the second regularized KE version, denoted by RKE-2 as follows.

Algorithm RKE-2. Let $x^0 \in I\!R^n, y^0 = b$; for $k = 0, 1 \dots$ do

$$y^{k+1} = \Phi(\alpha; y^k); \;\; b^{k+1} = b - y^{k+1}; \;\; x^{k+1} = F(\omega; b^{k+1}; x^k) - \gamma^2 Rx^k. \tag{15}$$

Remark 1 *Unfortunately, we have not yet a systematic convergence analysis for the above algorithm* RKE-2 *(although some "intuitive" arguments allow us to conjecture this). But, it has very good reconstruction properties as we have reported in the examples described in the next section.*

3 Numerical experiments

We present here our results for four image reconstruction experiments. We want to reconstruct the images from figures 1a (see [6]) and 2a. For each image we tested the previous algorithms with two initial approximations: $x_i^0 = 0$ (ie. zero initialization) and $x_i^0 = \bar{x}$ (i.e. Herman's [3] initialization), $\forall i = 1, \ldots, n$, where

$$\bar{x} = \frac{\sum_{i=1}^m b_i}{\sum_{i=1}^m \sum_{j=1}^n (A)_{ij}}.$$

In order to create the $n \times n$ symmetric and positive semidefinite matrix R we used the following method (see also [6]): for each $i \in \{1, \ldots, n\}$, let H_i be the set of horizontally neightbour pixels of i, V_i – the set of vertically neightbour pixels of i, and D_i – the set of diagonally neightbour pixels of i. Now, for each $j \in \{1, \ldots, n\}$,

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

where w_h, w_v, and w_d are parameters of the construction procedure.

The parameters we used for our tests are presented in table 1. For the construction of R we used $w_h = -1$, $w_v = -1$, $w_d = -1/\sqrt{2}$. With these values, the algorithms performed better than other values we tested (eg. $w_h = w_v = w_d = -1$, or $w_h = 1/3$, $w_v = 3$, $w_d = 1/\sqrt{1/3^2 + 3^2}$). We also tested RKE-1 with $\gamma = 10^{-2}$ and $\gamma = 5 \cdot 10^{-3}$, but the results were less satisfactory.

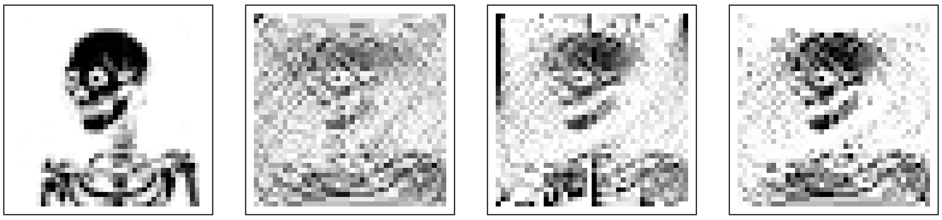
Parameter	**KE**	**RKE-1**	**RKE-2**
α	0.5	0.5	0.5
ω	0.8	0.8	0.8
γ	—	$5 \cdot 10^{-2}$	10^{-2}

Table 1: Algorithm settings for the tests

The reconstruction results for Herman's initialization are presented in figures 1 and 2, subfigures b), c), and d). Due to space limitation, we did not include reconstructed images for $x^0 = 0$, but we can say that the final results — when the maximum number of iterations was reached — were similar to those obtained with Herman's initialization. For the image

a) Original b) KE c) RKE-1 d) RKE-2

Figure 1: Test 1 image and reconstructions

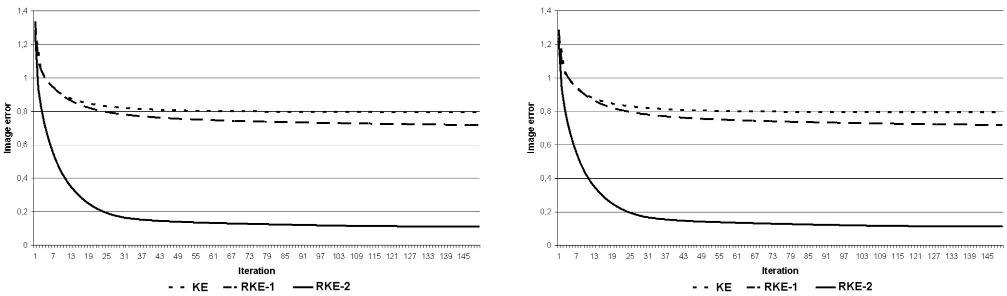


a) Original b) KE c) RKE-1 d) RKE-2

Figure 2: Test 2 image and reconstructions

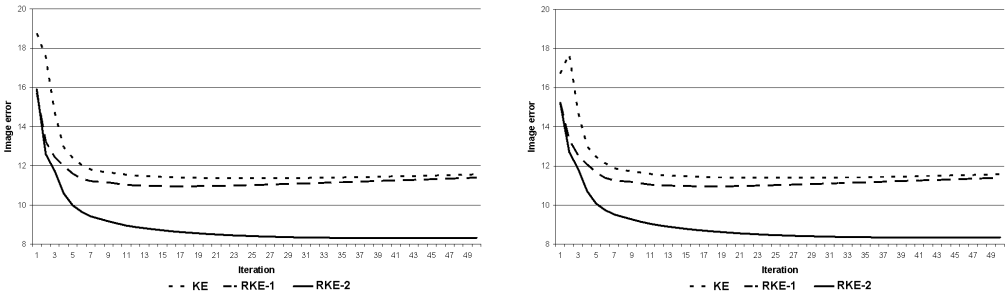
from figure 1a we allowed the algorithm to run 150 iterations, whereas for figure 2a only 50 iterations.

In order to evaluate the quality of the reconstructed image x relative to the original image x_{ex} we used the image error function $e_{x_{ex}}(x) = ||x - x_{ex}||$. Figures 3 and 4 show the evolution of the image errors for each iteration. As you may notice, the evolution of the image error differs between the two initialization methods during the first iterations, but the final errors are very close, thus the algorithm created similar reconstructions. Table 2 shows the image errors of the reconstructed images at the end of the experiments.



$x_i^0 = 0$ $x_i^0 = \bar{x}$

Figure 3: Test 1 reconstructions errors



$x_i^0 = 0$ $\qquad$ $x_i^0 = \bar{x}$

Figure 4: Test 2 reconstructions errors

Algorithm	Initialization	Reconstruction error	
		Test 1	Test 2
KE	Zero	0.7959	11.5644
KE	Herman	0.7949	11.5763
RKE-1	Zero	0.7187	11.3980
RKE-1	Herman	0.7187	11.3980
RKE-2	Zero	0.1125	8.3250
RKE-2	Herman	0.1127	8.3507

Table 2: Reconstruction errors

References

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

[2] *Engl, H. W., Hanke, M.* and *Neubauer, A.:* Regularization of inverse problems. Kluwer Academic Publishers, Dordrecht, 2000.

[3] *Herman, G.* and *Meyer, B. L.:* Algebraic reconstruction techniques can be made computationally efficient. IEEE Transactions on Medical Imaging, 12(3)(1993), 600-609.

[4] *Jiang, M.* and *Wang, G.:* Convergence studies on iterative algorithms for image reconstruction. IEEE Transactions on Medical Imaging, 22(5)(2003), 569-579.

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

[6] *Popa, C.* and *Zdunek, R.:* Penalized Least-Squares Image Reconstruction for Borehole Tomography, Proceedings of ALGORITMY 2005 Conference, Vysoke Tatry – Podbanske, Slovakia, March 13 – 18, 2005, 260–269.