

# A NOTE ON KACZMARZ ALGORITHM WITH REMOTEST SET CONTROL SEQUENCE \*

Constantin Popa<sup>†</sup>

Dedicated to Professor Mihail Megan  
on the occasion of his 70th anniversary

## Abstract

In this paper we analyse the Kaczmarz projection algorithm with remotest set control of projection indices. According to this procedure, at each iteration the projection index is one which gives the maximal absolute value of the corresponding residual. We prove that for under-determined full row rank systems and under some assumptions valid for problems arising in algebraic reconstruction of images in computerized tomography, this selection procedure has the property that each row index is selected at least once during the Kaczmarz algorithm iterations.

MSC: 65F10, 65F20

**keywords:** Kaczmarz algorithm; remotest set control

## 1 Introduction

For  $A$  an  $m \times n$  (real) matrix  $A$  and  $b \in \mathbb{R}^m$  in this paper we will consider the consistent (compatible) system of linear equations

$$Ax = b, \tag{1}$$

---

\* Accepted for publication on February 20, 2018

<sup>†</sup> [cpopa1956@gmail.com](mailto:cpopa1956@gmail.com); [cpopa@univ-ovidius.ro](mailto:cpopa@univ-ovidius.ro); Ovidius University of Constanta and Academy of Romanian Scientists, Romania

and denote by  $S(A; b)$  the set of its solutions and by  $x_{LS}$  the minimal (Euclidean) norm one. We will use the notations  $A^T, A_i, A^j, \mathcal{R}(A), \mathcal{N}(A), \text{rank}(A)$ , and  $P_V$  for the transpose,  $i$ -th row,  $j$ -th column, range and null space of  $A$ , the rank of  $A$ , and the projection onto a nonempty closed convex set  $V$ , respectively. We know that

$$\mathcal{R}(A^T) = \text{span}\{A_1, A_2, \dots, A_m\}, \quad x_{LS} \in \mathcal{R}(A^T).$$

Also  $\langle \cdot, \cdot \rangle$  and  $\| \cdot \|$  will denote the Euclidean scalar product and norm and all the vectors appearing in the paper will be considered as column vectors. If  $H_i = \{x \in \mathbb{R}^n, \langle x, A_i \rangle = b_i\}$  is the hyperplane determined by the  $i$ -th equation of the system (1) we have

$$P_{H_i}(x) = x - \frac{\langle x, A_i \rangle - b_i}{\|A_i\|^2} A_i. \quad (2)$$

The Kaczmarz algorithm with single projection (for short Kaczmarz) is the following.

**Algorithm Kaczmarz**

*Initialization:*  $x^0 \in \mathbb{R}^n$

*Iterative step:* for  $k = 0, 1, \dots$  select  $i_k \in \{1, 2, \dots, m\}$  and compute  $x^{k+1}$  as

$$x^{k+1} = x^k - \frac{\langle x^k, A_{i_k} \rangle - b_{i_k}}{\|A_{i_k}\|^2} A_{i_k}. \quad (3)$$

For an almost complete overview on the selection procedures in Kaczmarz algorithm see [2], [3] (section 5.1), [5], [4] and references therein. But, an important problem when considering a selection procedure seems to be the following: “*sooner or latter*” during the iterations each (row) projection index  $i_k$  must appear. This was clearly formulated in [4] as follows ( $\mathbb{N}$  will denote the set of natural numbers  $\{0, 1, 2, \dots, \}$ ).

**Definition 1** Given a monotonically increasing sequence  $\{\tau_k\}_{k=0}^\infty \subset \mathbb{N}$ , a mapping  $i : \mathbb{N} \rightarrow \{1, 2, \dots, m\}$  is called a control with respect to the sequence  $\{\tau_k\}_{k=0}^\infty$  if it defines a sequence  $\{i(t)\}_{t=0}^\infty$ , such that for all  $k \geq 0$ ,

$$\{1, 2, \dots, m\} \subseteq \{i(\tau_k), i(\tau_k + 1), \dots, i(\tau_{k+1} - 1)\}.$$

The set  $\tau_k, \tau_k + 1, \dots, \tau_{k+1} - 1$  is called the  $k$ -th window (with respect to the given sequence  $\{\tau_k\}_{k=0}^\infty$ ) and  $C_k = \tau_{k+1} - \tau_k$  its length. If the sequence of lengths  $(C_k)_{k \geq 0}$  is bounded the control  $\{i(t)\}_{t=0}^\infty$  itself is called bounded. If the sequence of lengths  $(C_k)_{k \geq 0}$  is unbounded the control  $\{i(t)\}_{t=0}^\infty$  itself is called an expanding control.

In the same paper [4] there are defined different types of bounded and expanding control sequences. But, there are also other types of control sequences which are not included in the above definition. Two well-known such examples are the *random control* and *remotest set control* (called in the present paper *Maximal Residual control* (MR, for short).

- **Maximal Residual control:** Select  $i_k \in \{1, 2, \dots, m\}$  such that

$$|\langle A_{i_k}, x^{k-1} \rangle - b_{i_k}| = \max_{1 \leq i \leq m} |\langle A_i, x^{k-1} \rangle - b_i|.$$

- **Random control:** Let the set  $\Delta_m \subset \mathbb{R}^m$  be defined by

$$\Delta_m = \{x \in \mathbb{R}^m, x \geq 0, \sum_{i=1}^m x_i = 1\},$$

define the discrete probability distribution

$$p \in \Delta_m, p_i = \frac{\|A_i\|^2}{\|A\|_F^2}, i = 1, \dots, m, \quad (4)$$

and select  $i_k \in \{1, 2, \dots, m\}$

$$i_k \sim p. \quad (5)$$

At least related to author's knowledge, there are no results saying that the above two control sequences satisfy the previously mentioned property, i.e. "sooner or latter" during the iterations of Kaczmarz algorithm (3) with that specific choice of the control sequence, each (row) projection index  $i_k$  must appear. More clear, we formulate this property as follows: "Determine appropriate assumptions on (1) such that

$$\forall i \in \{1, 2, \dots, m\}, \exists k \geq 0, \text{ with } i_k = i. \quad (6)$$

In the rest of the paper we will analyse this property for the Kaczmarz algorithm with Maximal Residual control sequence (MRK, for short) and show that it exists a case in which the property (6) can be theoretically proved.

## 2 Algorithm MRK

We consider in this section Kaczmarz algorithm (3) in which the Maximal Residual control procedure is used for selecting the projection indices at each iteration (caled MRK algorithm).

**Algorithm MRK**

*Initialization.*  $x^0 \in \mathbb{R}^n$ ;

*Iterative step.* Select  $i_k \in \{1, \dots, m\}$  such that

$$|r_{i_k}| = \max_{1 \leq i \leq m} |r_i|, \text{ where } r = Ax^k - b \in \mathbb{R}^m,$$

and perform the projection

$$x^{k+1} = P_{H_{i_k}}(x^k), \forall k \geq 0.$$

The following result gives us a sufficient condition such that the property (6) holds.

**Proposition 1** *Let  $m \leq n$  and suppose that*

$$\text{rank}(A) = m, \tag{7}$$

and

$$P_{\mathcal{R}(A^T)}(x^0) - x_{LS} = \sum_{i=1}^m \gamma_i A_i.$$

If

$$\gamma_i \neq 0, \forall i = 1, \dots, m, \tag{8}$$

then (6) is true for the MRK algorithm.

**Proof.** Suppose that (6) is not satisfied and let  $i_0 \in \{1, 2, \dots, m\}$  be such that  $i_k \neq i_0$ , for all  $k \geq 0$ . Then, (2) yields that

$$x^k = x^0 + \sum_{1 \leq i \leq m; i \neq i_0} \alpha_i^k A_i,$$

with  $\alpha_i^k \in \mathbb{R}$ , hence

$$x^k \in x^0 + \text{span}(A_i, i = 1, \dots, m, i \neq i_0). \tag{9}$$

In [1] the author proved that for consistent systems as (1) (which holds in our case because of the assumption (7)) the sequence  $(x^k)_{k \geq 0}$  generated with the MRK algorithm converges and

$$\lim_{k \rightarrow \infty} x^k = P_{\mathcal{N}(A)}(x^0) + x_{LS}. \tag{10}$$

Since the set  $x^0 + \text{span}(A_i, i = 1, \dots, m, i \neq i_0)$  is closed, from (9)) it results that the limit vector in (10)) belongs to the same set, thus

$$P_{\mathcal{N}(A)}(x^0) + x_{LS} - x^0 = x_{LS} - P_{\mathcal{R}(A^T)}(x^0) \in \text{span}(A_i, i = 1, \dots, m, i \neq i_0).$$

This contradicts the hypothesis (8) and completes the proof.  $\spadesuit$

Regarding the possibility to fulfil the hypothesis (8), we give the following result.

**Proposition 2** *Let*

$$x^0 = \sum_{i=1}^m \beta_i A_i \in \mathcal{R}(A^T)$$

and suppose that

$$A_i \neq 0, A_{ij} \geq 0, \forall i, j \text{ and } \|x_{LS}\| \leq M, \quad (11)$$

for some  $M \geq 0$ . If the scalars  $\beta_i$  satisfy

$$\beta_i > \frac{M}{M_i}, M_i = \max_{1 \leq j \leq n} A_{ij} > 0, \forall i = 1, \dots, m, \quad (12)$$

then

$$P_{\mathcal{R}(A^T)}(x^0) - x_{LS} = x^0 - x_{LS} = \sum_{i=1}^m \gamma_i A_i, \text{ with } \gamma_i > 0, \forall i = 1, \dots, m. \quad (13)$$

**Proof.** Let

$$x_{LS} = (x_1, \dots, x_n)^T = \sum_{i=1}^m \alpha_i A_i. \quad (14)$$

We distinguish the following two cases.

**Case 1.** Let  $i_0 \in \{1, \dots, m\}$  be an index such that in (14)  $\alpha_{i_0} \leq 0$ . Then, if we take  $\beta_{i_0} > 0$  for the corresponding  $\gamma_{i_0}$  in (13) we obtain  $\gamma_{i_0} > 0$ , which fits into our conclusion.

**Case 2.** According to **Case 1** we may suppose that in (14) we have

$$\alpha_i > 0, \forall i = 1, \dots, m.$$

From (14) we get

$$0 \leq x_j = \sum_{i=1}^m \alpha_i A_{ij}, \forall j = 1, \dots, n,$$

which gives us

$$x_j = |x_j| \leq \sqrt{\sum_{q=1}^n x_q^2} = \|x_{LS}\| \leq M, \quad \forall j = 1, \dots, n,$$

and therefore

$$0 \leq \sum_{i=1}^m \alpha_i A_{ij} \leq M, \quad \forall j = 1, \dots, n. \quad (15)$$

If  $i \in \{1, \dots, m\}$  is arbitrary fixed, from (15) we obtain

$$0 \leq \alpha_i A_{ij} \leq M, \quad \forall j = 1, \dots, n. \quad (16)$$

Again because of our assumptions (11) it results that it exists at least one index  $j$  such that  $A_{ij} > 0$ , which tell us that

$$M_i = \max_{1 \leq j \leq n} A_{ij} > 0. \quad (17)$$

From (15) - (16) we obtain that the coefficients  $\alpha_i$  from (14) should satisfy

$$\alpha_i \leq \frac{M}{M_i}, \quad \forall i = 1, \dots, m$$

with  $M_i$  defined in (17). Hence, in order to get the conclusion (13) we must take  $\beta_i$  as in (12) and the proof is complete.  $\spadesuit$

**Remark 1** *If  $A$  is a scanning matrix from ART in CT, the second assumption in (11) is satisfied. The first assumption is usually imposed for projection-based iterative methods. Anyway, it is not a restrictive condition because any zero row from  $A$  can be eliminated from the beginning without changing the solution set of (1). The third assumption is also connected with the ART; indeed we usually have information about the components of the solutions  $z = (z_1, \dots, z_n)^T \in S(A; b)$  of the form  $0 \leq z_j \leq C, \forall j = 1, \dots, n$ . This gives us*

$$\|x_{LS}\| \leq \|z\| = \sqrt{\sum_{j=1}^n z_j^2} \leq \sqrt{n}C.$$

**Remark 2** *In the case of random control (4) - (5), the sequence  $(x^k)_{k \geq 0}$  generated by the corresponding Kaczmarz iteration (3) starts with  $x^0 = 0$  and has the property (see e.g. [6], [7])*

$$\lim_{k \rightarrow \infty} x^k = x_{LS}.$$

Hence, in this case the assumption (8) would become

$$x_{LS} = \sum_{i=1}^m \gamma_i A_i, \text{ with } \gamma_i \neq 0, \forall i = 1, \dots, m$$

which is hard to be verified.

## References

- [1] Ansorge R., *Connections between the Cimmino-method and the Kaczmarz-method for the solution of singular and regular systems of equations*, Computing, **33** (1984), 367-375.
- [2] Censor, Y., *Row-action methods for huge and sparse systems and their applications*, SIAM Review, 23 (1981), 444-466.
- [3] Censor, Y., Zenios, S., *Parallel Optimization: Theory, Algorithms and Applications*, Oxford Univ. Press (1997)
- [4] Censor, Y. Chen, W., Pajooohesh, H., *Finite convergence of a subgradient projections method with expanding controls*, Appl. Math. Optim., **64** (2011), 273–285.
- [5] Combettes P., *Hilbertian convex feasibility problem: Convergence of projection methods*, Appl. Math. Optim., **35** (1997), 311–330.
- [6] Strohmer, T., Vershynin, R., *A randomized Kaczmarz algorithm with exponential convergence*, J. Fourier Anal. Appl. 15(2009), 262 - 278.
- [7] Zouzias, A., Freris, N.: *Randomized Extended Kaczmarz for Solving Least Squares*. SIAM J. Matrix Anal. Appl. 34(2), 773 - 793 (2013)