DOI: 10.2478/auom-2018-0015
An. Şt. Univ. Ovidius Constanţa **Vol. 26(1),2018, 241–247**

On Remotest set and Random controls in Kaczmarz algorithm

Constantin Popa

Abstract

In this paper we analyse the Kaczmarz projection algorithm with Remotest set and Random control of projection indices and provide a sufficient condition such that each projection index appears infinitely many times during the iterations.

1 Introduction

For an $m \times n$ (real) matrix A and $b \in I\!R^m$ let

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

be a consistent system of linear equations and denote by $S(A; b), x_{LS}$ the set of its solutions and the minimal (Euclidean) norm one ($\langle \cdot, \cdot \rangle$ and $\| \cdot \|$ will denote the Euclidean scalar product and norm on some space $I\!R^q$, respectively). Other notations used will be $A^T, A_i, A^j, \mathcal{R}(A), \mathcal{N}(A)$, $rank(A)$ for the transpose, i-th row, j-th column, range, null space and rank of A. The projection onto a nonempty closed convex set V will be denoted by P_V, and for $V = H_i = \{x \in I\!R^n, \langle x, A_i \rangle = b_i\}$ (the hyperplane determined by the i-th equation of the system (1)) we know that

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

Key Words: Kaczmarz algorithm; control sequence; remotest set control; random control.
2010 Mathematics Subject Classification: 65F10, 65F20.
Received: 20.12.2016
Revised: 10.04.2017
Accepted: 20.04.2017

The Kaczmarz's iterative method for numerical solution of (1) has the form from below.

Algorithm K.

Initialization: $x^0 \in I\!R^n$

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

$$x^{k+1} = P_{H_{i_k}}(x^k). \tag{3}$$

There have been defined several classes of selection procedures for the indices i_k (see [4, 3, 5, 6] and references therein). In this paper we will consider the Maximal Residual (remotest set) control and the Random control procedures, and provide a sufficient condition such that in the case of Kaczmarz's projection method **K**, they belong to the class of *control* selections from [5]. The paper is organized as follows: in section 2 we give an equivalent formulation of the *control* sequence definition from [5]. In section 3 we show that, for $x^0 = 0$ and under additional assumptions, the Maximal Residual (remotest set) selection or the Random selection is a *control* w.r.t. section 2.

2 Control sequences

Let $I\!N$ denote the set of natural numbers $\{0, 1, 2, \dots, \}$. In [5] the following definition concerning control sequences was introduced.

Definition 1. (D1) *Given a monotonically increasing sequence* $\{\tau_k\}_{k\geq 0} \subset I\!N$*, a mapping* $i : I\!N \to \{1, 2, \dots, m\}$ *is called a* control with respect to the sequence $\{\tau_k\}_{k\geq 0}$ *if it defines a* sequence $\{i(t)\}_{t\geq 0}$*, such that for all* $k \geq 0$*,*

$$\{1, 2, \dots, m\} \subseteq \{i(\tau_k), i(\tau_k + 1), \dots, i(\tau_{k+1} - 1)\}. \tag{4}$$

The next definiton, mentioned in [2] (see also [7]) points-out on an important aspect of control sequences.

Definition 2. (D2) *A mapping* $i : I\!N \to \{1, 2, \dots, m\}$ *is called a* random mapping *if any* $i \in \{1, \dots, m\}$ *appears infinitely many times in the set* $\mathcal{I} = \{i(k), k \geq 0\}$*.*

It is clear that, if the mapping i is a control with respect to some sequence $\{\tau_k\}_{k\geq 0}$, then it is also a random mapping, according to definition (**D2**). Indeed, if the sequence $\{\tau_k\}_{k\geq 0}$ is increassing then $\tau_{k+1} > \tau_k$ and the sets $\Delta_k = \{i(\tau_k), i(\tau_k + 1), \dots, i(\tau_{k+1} - 1)\}, k \geq 0$ form a partition of $I\!N$ as in (4). Next proposition tell us about the reciprocal of this property, i.e. a random mapping is a control according to the definition (**D1**).

Proposition 1. *Let* i : $I\!N \to \{1, 2, \ldots, m\}$ *be a random mapping (according to* **(D2)***). Then it exists a monotonically increasing sequence* $\{\tau_k\}_{k\geq 0} \subset I\!N$ *such that* i *is a control w.r.t.* **(D1)**.

Proof. We will first write **(D2)** in the following equivalent formulation: for any $i \in \{1, \ldots, m\}$ it is true that

$$\forall\ k \geq 1,\ \exists k_i \geq k \text{ s.t. } i_k = i. \tag{5}$$

We will now recursively define an increasing sequence $\{\tau_k\}_{k\geq 0} \subset I\!N$ as follows: for $k = 0$ we set $\tau_0 = 0$; for $k = 1$ let τ_1 be the smallest natural number with the properties

$$\tau_1 > \tau_0 \text{ and } \{1, 2, \ldots, m\} \subseteq \{i(\tau_0), \ldots, i(\tau_1 - 1)\}. \tag{6}$$

Such a number τ_1 exists according to the equivalent formulation (5). In general, if we already have constructed τ_k, then τ_{k+1} will be the smallest natural number such that

$$\tau_{k+1} > \tau_k \text{ and } \{1, 2, \ldots, m\} \subseteq \{i(\tau_k), \ldots, i(\tau_{k+1} - 1)\}, \tag{7}$$

which is exactly the property (4) of definition **(D1)**, and the proof is complete. □

Based on the above proposition we will consider in the rest of the paper as definition for controls the equivalent formulation from **(D2)**. In this respect, the following two selection procedures will be analysed.

- **Maximal Residual (remotest set) ([1]):** Select $i_k \in \{1, 2, \ldots, m\}$ such that

$$|\langle A_{i_k}, x^{k-1}\rangle - b_{i_k}| = \max_{1\leq i\leq m} |\langle A_i, x^{k-1}\rangle - b_i|. \tag{8}$$

- **Random ([9]):** Let the set $\Delta_m \subset I\!R^m$ be defined by

$$\Delta_m = \{x \in I\!R^m, x \geq 0, \sum_{i=1}^{m} = 1\}, \tag{9}$$

  define the discrete probability distribution

$$p \in \Delta_m,\ p_i = \frac{\|A_i\|^2}{\|A\|_F^2},\ i = 1, \ldots, m, \tag{10}$$

  and select $i_k \in \{1, 2, \ldots, m\}$ such that

$$i_k \sim p. \tag{11}$$

The main aspect regarding the above two selection procedures is concerned with the fact that the projection indices are generated recursively, without no a priori information on them. And, at least related to author's knowledge, there are no results saying that when the algorithm **K** is applied with one or the other of the above selection procedures, each projection index will appear infinitely many times.

3 The Kaczmarz algorithm

We consider in this section Kaczmarz's projection algorithm in which the Maximal Residual (remotest set) (8) or Random (10)-(11) procedure is used for selecting the projection indices in each iteration, and with the initial approximation $x^0 = 0$. In this case, in papers [1] and [9] it is proved that the sequence $(x^k)_{k\geq 0}$ generated by algorithm **K** converges to the minimal norm solution x_{LS} of the system (1). We will formulate a sufficient condition such that any of the above selection procedures satisfies **(D2)**). For $i \in \{1, \ldots, m\}$ arbitrary fixed, let $A^{(i)} : (m-1) \times n$, $b^{(i)} \in I\!R^{m-1}$ be the submatrix of A without the i-th row, respectively the subvector of b without the i-th component and $x_{LS}^{(i)}$ the minimal norm solution of the system $A^{(i)}x = b^{(i)}$.
Assumption C. For any index $i \in \{1, \ldots, m\}$ we have

$$x_{LS} \neq x_{LS}^{(i)}. \tag{12}$$

Proposition 2. *If the assumption* **C** *holds, then any of the above two selection procedures within the Kaczmarz's iteration* **K** *satisfies* **(D2)**.

Proof. Let us suppose that the conclusion of the proposition is not true. According to (5) it exists an index $i_0 \in \{1, \ldots, m\}$ and an integer $k_0 \geq 1$ such that, in the selection procedure of the **K** algorithm iterations we have

$$i_k \neq i_0, \ \forall k \geq k_0. \tag{13}$$

Therefore, the sequence $(x^k)_{k\geq k_0}$ is generated by the **K** algorithm applied (only !) to the subsystem $A^{(i_0)}x = b^{(i_0)}$. By the theory from [1] and [9], respectively, it results that

$$\lim_{0\geq k\to\infty} x^k = x_{LS} = \lim_{k_0\geq k\to\infty} x^k = x_{LS}^{(i)}, \tag{14}$$

hence

$$x_{LS} = x_{LS}^{(i_0)}, \tag{15}$$

which contradicts (12) and completes the proof. $\square$

In order to understand what means a condition like (12) we will analyse it in the particular case

$$m \leq n \text{ and } rank(A) = m, \tag{16}$$

for which the system (1) is consistent for any $b \in \mathbb{R}^m$. In this case for any index i the matrix $A^{(i)}$ is also full-row rank, and the system $A^{(i_0)}x = b^{(i_0)}$ also consistent. We will arbitrary fix the index $i \in \{1, \ldots, m\}$ and denote by $\tilde{A}, \tilde{b}, \tilde{x}_{LS}$ the elements $A^{(i)}, b^{(i)}, x_{LS}^{(i)}$, respectively. Moreover, we will analyse the opposite assumption of (12), namely

$$x_{LS} \ = \ \tilde{x}_{LS}. \tag{17}$$

What does this mean in terms of the matrix A and right hand side b ? For simplfying the presentation we will suppose that $i = m$ (this assumption is not too restrictive because it can be obtained by a row-permutation in A and b, which does not affect the spectral properties of A and the solution set $S(A; b)$). Because A^T is overdetermined and full-column rank, there exist an $n \times n$ orthogonal matrix Q and the QR decomposition

$$Q^T A^T = \begin{bmatrix} R \\ 0 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & \ldots & r_{1,m-1} & r_{1m} \\ 0 & r_{22} & \ldots & r_{2,m-1} & r_{2m} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & r_{m-1,m-1} & r_{m-1,m} \\ 0 & 0 & \ldots & 0 & r_{mm} \\ 0 & 0 & \ldots & 0 & 0 \\ \vdots & & & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 0 \end{bmatrix} =$$

$$Q^T [\tilde{A}^T | A_m] = [Q^T \tilde{A}^T | Q^T A_m] = \begin{bmatrix} \tilde{R} & & c \\ 0 & & r_{mm} \\ 0 & \ldots & 0 \\ \vdots & & \vdots \\ 0 & \ldots & 0 \end{bmatrix}. \tag{18}$$

Therefore

$$Q^T \tilde{A}^T = \begin{bmatrix} \tilde{R} \\ 0 \end{bmatrix} \tag{19}$$

will be a QR decomposition for $\tilde{A}$, where

$$\tilde{R} = \begin{bmatrix} r_{11} & r_{12} & \ldots & r_{1,m-1} \\ 0 & r_{22} & \ldots & r_{2,m-1} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & r_{m-1,m-1} \end{bmatrix} \text{ and } c = (r_{1m}, r_{2m}, \ldots, r_{m-1,m})^T. \tag{20}$$

Because $m \leq n$ and $A, \tilde{A}$ have full-row rank, we know that (see e.g. [8])

$$A^+ = A^T(AA^T)^{-1}, \quad \tilde{A}^+ = \tilde{A}^T(\tilde{A}\tilde{A}^T)^{-1},$$

hence

$$x_{LS} = A^+ b = Q \begin{bmatrix} R^{-T} b \\ 0 \end{bmatrix}, \quad \tilde{x}_{LS} = \tilde{A}^+ \tilde{b} = Q \begin{bmatrix} \tilde{R}^{-T} \tilde{b} \\ 0 \end{bmatrix}. \tag{21}$$

In our hypothesis (17), and by using (18) and (20) we get from (21) the equality

$$R^{-T} b = \begin{bmatrix} \tilde{R}^{-T} \tilde{b} \\ 0 \end{bmatrix}, \text{ where } R^T = \begin{bmatrix} \tilde{R}^T & 0 \\ c^T & r_{mm} \end{bmatrix} : m \times m. \tag{22}$$

It can be easily shown that

$$R^{-T} = (R^T)^{-1} = \begin{bmatrix} \tilde{R}^{-T} & 0 \\ -\frac{1}{r_{mm}} c^T \tilde{R}^{-T} & \frac{1}{r_{mm}} \end{bmatrix}, \tag{23}$$

which together with the first equality in (22) gives

$$\begin{bmatrix} \tilde{R}^{-T} \tilde{b} \\ 0 \end{bmatrix} = \begin{bmatrix} \tilde{R}^{-T} & 0 \\ -\frac{1}{r_{mm}} c^T \tilde{R}^{-T} & \frac{1}{r_{mm}} \end{bmatrix} \begin{bmatrix} \tilde{b} \\ b_m \end{bmatrix} = \begin{bmatrix} \tilde{R}^{-T} \tilde{b} \\ -\frac{1}{r_{mm}} c^T \tilde{R}^{-T} \tilde{b} + \frac{b_m}{r_{mm}} \end{bmatrix},$$

and therefore

$$0 \; = \; -\frac{1}{r_{mm}} c^T \tilde{R}^{-T} \tilde{b} + \frac{b_m}{r_{mm}} \text{ or } c^T \tilde{R}^{-T} \tilde{b} = b_m.$$

Eventually, we proved that if (17) holds (for $i = m$), then

$$c^T \tilde{R}^{-T} \tilde{b} = b_m, \tag{24}$$

where the elements $c, \tilde{R}$ are from (18) and $\tilde{b}$ is the right hand side of the system $A^{(i_0)} x = b^{(i_0)}$. But, also the converse holds, namely: if (24) is true with the above elements, then (17) holds (for $i = m$). This is true if we assume that in the QR decomposition (18) the diagonal elements satisfy $r_{ii} > 0, \forall i$, which gives us the unicity of the factor R in the QR decomposition.

Remark 1. *Although the assumption $x^0 = 0$ in $\mathbf{K}$ is essential for the proof of Proposition 2, we conjecture that this result is stil true for a larger class of initial approximations x^0. Unfortunately we do not have for the moment a theoretical proof in this respect.*

Acknowledgements. We would like to thanks to Prof. Yair Censor for his very helpful comments that have much improved the initial version of the paper.

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] Bauschke H.H., Borwein J.M., *Legendre functions and the method of random Bregman projections*, J. Convex Anal., 4(1997), 27-67.

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

[4] Censor, Y., Zenios, S., *Parallel Optimization: Theory, Algorithms and Applications*, Oxford Univ. Press (1997)

[5] Censor, Y. Chen, W., Pajoohesh, H., *Finite convergence of a subgradient projections method with expanding controls*, Appl. Math. Optim., **64** (2011), 273–285.

[6] Combettes P., *Hilbertian convex feasibility problem: Convergence of projection methods*, Appl. Math. Optim., **35** (1997), 311–330.

[7] Lorenz, D.A. et al., *The linearized Bregman Method via Split Feasibility Problems: Analysis and Generalizations*, SIAM J. Imaging Sciences, 7(2)(2014), 1237-1262.

[8] Popa C. - *Projection algorithms - classical results and developments. Applications to image reconstruction*, Lambert Academic Publishing - AV Akademikerverlag GmbH & Co. KG, Saarbrücken, Germany, 2012

[9] Strohmer, T., Vershynin, R., *A randomized Kaczmarz algorithm with exponential convergence*, J. Fourier Anal. Appl. 15(2009), 262 - 278.

[10] Zouzias A., Freris N., *Randomized Extended Kaczmarz for Solving Least Squares*, SIAM Journal on Matrix Analysis and Applications, 34(2)92013), 773-793.

Constantin POPA,
Ovidius University of Constanta,
Mamaia 124, 900527 Constanta, Romania.

“Gheorghe Mihoc - Caius Iacob” Institute of
Statistical Mathematics and Applied Mathematics,
Calea 13 Septembrie 13, 050711 Bucharest, Romania.

Academy of Romanian Scientists,
Splaiul Independentei 54, 050085 Bucharest, Romania
Email: cpopa@univ-ovidius.ro