TECHNICAL TRANSACTIONS | **CZASOPISMO TECHNICZNE**
FUNDAMENTAL SCIENCES | NAUKI PODSTAWOWE

2-NP/2015

ROMAN SZNAJDER*

KACZMARZ ALGORITHM REVISITED

JESZCZE O ALGORYTMIE KACZMARZA

A b s t r a c t

In 1937, Stefan Kaczmarz proposed a simple method, called the *Kaczmarz algorithm*, to solve iteratively systems of linear equations $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ in Euclidean spaces. This procedure employs cyclic orthogonal projections onto the hyperplanes associated with such a system. In the case of a nonsingular matrix $\mathbf{A}$, Kaczmarz showed that his method guarantees convergence to the solution of $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$. The Kaczmarz algorithm was rediscovered in 1948 and became an important tool in medical engineering. We briefly discuss generalizations of this method and its ramifications, including applications in computer tomography, image processing and contemporary harmonic analysis.

Keywords: *Kaczmarz method*, *systems of linear equations*, *computer tomography*, *image reconstruction*

S t r e s z c z e n i e

W 1937 roku Stefan Kaczmarz zaproponował prostą metodę [*KA*], zwaną obecnie *algorytmem Kaczmarza*, za pomocą której można rozwiązywać iteracyjnie układy równań liniowych $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ w przestrzeniach euklidesowych. Metoda ta używa cyklicznego ciągu rzutów ortogonalnych na hiperpłaszczyzny związane z tym układem. W przypadku macierzy odwracalnej $\mathbf{A}$ Kaczmarz pokazał, że jego metoda gwarantuje zbieżność do rozwiązania układu równań $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$. Metoda ta została ponownie odkryta w 1948 roku i stała się ważnym narzędziem w inżynierii medycznej. Omawiamy tutaj pokrótce uogólnienia tej metody i ich zastosowania w tomografii komputerowej, przetwarzaniu obrazów i we współczesnej analizie harmonicznej.

Słowa kluczowe: *metoda Kaczmarza*, *układy równań liniowych*, *tomografia komputerowa*, *rekonstrukcja obrazu*

DOI: 10.4467/2353737XCT.15.220.4425

* Department of Mathematics, Bowie State University, USA; rsznajder@bowiestate.edu

1. The origins of the Kaczmarz algorithm

Stefan Kaczmarz has been primarily known as a member of the Lvov School of Mathematics and a collaborator of Stefan Banach and Hugo Steinhaus (see a celebrated monograph [20])[1]. His professional interests were orthogonal series, theory of real functions, and applications of mathematics. The aim of this article is to highlight the importance of Kaczmarz's pioneering work [21], which has found numerous applications across a number of fields, including image processing, computer tomography, and image compression. The paper appeared in 1937 in German and is hardly known to the majority of the Polish mathematical community, although it became widely recognized in the Western hemisphere. For years, researchers have been using the German original, as the English version appeared only in 1993, translated by professor P.C. Parks [22].

Fig. 1. Stefan Kaczmarz (1895–1939)

It seems that the only note on Kaczmarz algorithm in Polish literature was by Cegielski [7]. In this paper, the author, a noted expert on modern iterative computational methods, refers to an extensive list [8] of English-language publications on the Kaczmarz method. In his recent monograph [9], among others, Cegielski studies convergence of such a type of methods.

In its original formulation, the Kaczmarz algorithm (*KA*) states the following: for a given $m \times n$ matrix $\mathbf{A}$ and a vector $\boldsymbol{b} \in R^m$, we wish to find a solution to the linear system $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$. Let $\boldsymbol{x}^0 \in R^n$. Define the sequence of vectors

$$\boldsymbol{x}^{k+1} := \mathbf{x}^k - \frac{< \boldsymbol{a}^i, \boldsymbol{x}^k > - b_i}{\left\| \boldsymbol{a}^i \right\|^2} \boldsymbol{a}^i, \quad \text{where } k + 1 \equiv i \ (\mathrm{mod}\ n)$$

[1] For a superb and exhaustive monograph on Kaczmarz's research and private life, see [25].

and $\boldsymbol{a}^i$ is the i-th row of matrix $\mathbf{A}$ (where $\|\cdot\|$ stands for the Euclidean norm in R^n). Kaczmarz originally considered systems with a square matrix and showed that for a nonsingular matrix $\mathbf{A}$, the sequence $(\boldsymbol{x}^k)$ converges to the solution, regardless of the starting point $\boldsymbol{x}^0 \in R^n$.

How does the algorithm work? For any $1 \leq k \leq n$, the above formula presents the orthogonal projection of the point $\boldsymbol{x}^k$ onto the affine hyperplane

$$H_i = \{\boldsymbol{x} \in R^n : <\boldsymbol{a}^i, \boldsymbol{x}> = b_i\} \quad (i = 1, \ldots, n).$$

The point $\boldsymbol{x}^{n+1}$ is projected again onto the hyperplane H_1, $\boldsymbol{x}^{n+2}$ is projected onto H_2, and so on. It is also called a *cyclic projection method*. In general, when A is of full rank, one gets $\boldsymbol{x}^k \rightarrow \boldsymbol{x}$ for some solution $\boldsymbol{x}$ to the system $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$.

The *KA* is historically the first numerical method exploring sequences of orthogonal projections onto hyperplanes. This method, while completely elementary – high school students should be able to grasp it – is quite powerful. On the other hand, certain issues such as the speed of convergence, were hard to settle.

In the simplest case of two intersecting lines l_1 and l_2 on the plane with normal (linearly independent) vectors $\boldsymbol{a}^1$ and $\boldsymbol{a}^2$, for a given seed point $\boldsymbol{x}^0$, $(\boldsymbol{x}^k)$ is the sequence of alternating orthogonal projections on these lines. The sequence obviously converges to the common point $\boldsymbol{x}$ of the given lines.

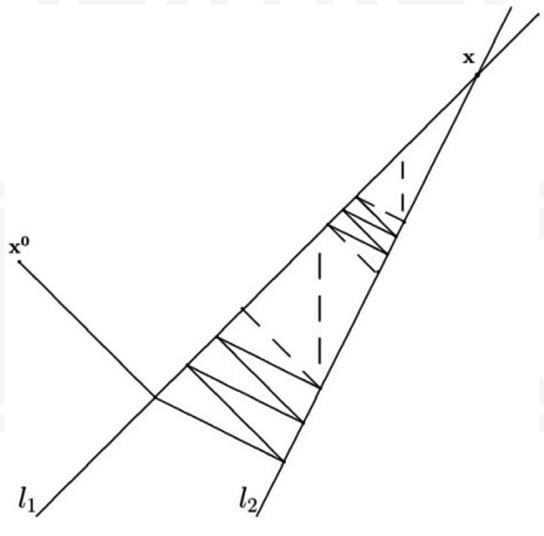

Fig. 2. The Kaczmarz method in the case of intersecting lines on the plane

Soon after the Kaczmarz's discovery, in 1938, Gianfranco Cimmino [10] proposed a similar iterative method: one reflects a given point $\boldsymbol{x}^k \in R^n$ about all hyperplanes H_i and averages these reflections with respect to a fixed probability vector $\boldsymbol{w} = (w_1, \ldots, w_n)$:

$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + 2\sum_{i=1}^{n} w_i (P_{H_i}\boldsymbol{x}^k - \boldsymbol{x}^k).$$

Cimmino showed that his method of simultaneous reflections converges to a solution. Both Kaczmarz and Cimmino algorithms were established ahead of their time and their invention passed in math community virtually unnoticed. These methods evolved from original forms and their generalized versions are currently widely used to solve large systems of linear equations. Although Cimmino's algorithm is slightly slower than the *KA*, the advantage of this method is the opportunity of using parallel processors. We should point out that Cimmino's method was extended to simultaneous projections onto closed convex sets.

2. Reemergence of the Kaczmarz algorithm

For more than a decade the *KA* was in oblivion. It resurfaced in separate publications of Bodewig [2] (1948), Forsythe [14] (1953), and Tompkins [32] (1949). The methods of projections were already known to these authors and the *KA* was performing remarkably well, despite slow computers in the early 1950s.

Systematic studies and applications of the *KA* started in 1970 with the paper by Gordon, Bender and Herman [15]. The *KA* was rediscovered by these authors and is known as the *Algebraic Reconstruction Technique* in computer tomography. For matrix $\mathbf{A} \in R^{m \times n}$, solvability of the matrix equation $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ indeed means finding (reconstructing) $\boldsymbol{x}$ from the data: $< \boldsymbol{a}^i, \boldsymbol{x} > = b_i$, $i = 1, 2, \ldots, m$; we assume here that $m \gg n$ (overdetermined system) and $\mathbf{A}$ is of full rank.

In 1971, Tanabe [31], undertook the effort of generalizing the *KA* and providing a deeper insight into the theoretical aspects of this method. He showed that the sequence $(\boldsymbol{x}^k)$ of iterates always converges, regardless of the consistency of the system $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$. He also noticed that the *KA* can be used to approximate the Moore-Penrose pseudoinverse $\mathbf{A}^\dagger$ of a matrix $\mathbf{A}$.

The *KA* was implemented in the first medical scanner in 1972. It became clear that there were a lot of potential applications of the *KA*. During the second half of the past century, the computational mathematics community witnessed an outburst of various iterative techniques, including simultaneous projections, relaxation and averaging techniques. Faster hardware and efficient software contributed to a rapid increase in applications of these techniques in medical sciences, e.g., in reconstruction of 3*D* images through 2*D* projections (computer tomography), see Brooks [4]. The Kaczmarz method can be considered as a special case of the POSC (*Projection onto Convex Sets*) method, see [9] and references therein, including Kiwiel [23], and the important survey paper by Bauschke and Borwein [1]. This technique plays a prominent role in signal and image processing, particularly in medical image processing, and in such disciplines as operations research and game theory. Byrne's monograph [6] emphasizes the importance of the so-called block Kaczmarz method; see also [5].

3. Evolution of the Kaczmarz method

Over the time, the *KA* evolved to become faster and more efficient. One of important modifications of the original *KA* is the so-called *randomized Kaczmarz algorithm*:

$$\boldsymbol{x}^{(k+1)} := \boldsymbol{x}^k + \frac{\boldsymbol{b}_{p(i)} - <\boldsymbol{a}^{p(i)}, \boldsymbol{x}^k>}{\left\| \boldsymbol{a}^{p(i)} \right\|^2} \boldsymbol{a}^{p(i)},$$

where, with $p(i) \in \{1, \ldots, m\}$, the row $\boldsymbol{a}^{p(i)}$ is chosen at random with probability $\frac{\left\| \boldsymbol{a}^{p(i)} \right\|^2}{\| \mathbf{A} \|_F^2}$; here, $\| \mathbf{A} \|_F$ is the Frobenius norm of matrix **A**, i.e., $\| \mathbf{A} \|_F^2 = \sum_{i=1}^m \left\| \boldsymbol{a}^i \right\|^2 =$ $= \sum_{i=1}^m \sum_{j=1}^n a_{ij}^2$. The magnitude of rows $\boldsymbol{a}^i$ of **A** cannot be ignored. Theoretical results on exponential convergence in expectation were obtained by Strohmer and Vershynin in their seminal work [30]. In fact, it is the first paper where such an estimate was obtained for the randomized *KA*. An interesting follow-up was done by Deanna Needell in [27]; see also her other work [18], [28], and [29]. In [13] the authors presented a modified version of the randomized *KA*, which in most cases significantly improved the convergence rate. They utilized the Johnson-Lindenstrauss dimension reduction technique to keep the runtime at the same order as that of original randomized version. The Johnson-Lindenstrauss lemma is a very interesting fact in its own right: it is one of the most quoted results in analysis, and deserves a separate exposition:

Let $\delta > 0$ *and S be a finite set of points in* R^n. *Then for any d satisfying*

$$d \geq C \frac{\log |S|}{\delta^2},$$

there exists a Lipschitz mapping $F : R^n \to R^d$ *such that*

$$(1-\delta) \left\| s_i - s_j \right\|^2 \leq \left\| F(s_i) - F(s_j) \right\|^2 \leq (1+\delta) \left\| s_i - s_j \right\|^2,$$

for all $s_i, s_j \in S$ [19].

In recent times, several master's and doctoral theses on spectral tomography and the block *KA* appeared, see [3, 4] (mentioned earlier), and [12]. An earlier dissertation by Grangeat [16] addressed the 3*D* image reconstruction from 2*D* X-ray pictures, however it did not directly refer to the *KA*. It is worthwhile to mention [33], which presented a practical application of the *KA*.

4. Extension of the Kaczmarz method to Hilbert spaces

In 1977, McCormick [26] was the first to investigate the *KA* in Hilbert spaces. In 2001, Kwapień and Mycielski [24] proposed an efficient generalization of the *KA* to infinite-dimensional Hilbert space. Here is their generalization: Let *H* be a Hilbert space and let $(\boldsymbol{e}^n)_{n=0}^{\infty}$ be a sequence of unit vectors in *H*. Given $\boldsymbol{x} \in H$, the Kaczmarz algorithm is defined as:

$$x^0 := < x, e^0 > e^0, \text{ and } \forall n \in N,\ x^n := x^{n-1} + < x - x^{n-1}, e^n > e^n.$$

It is important to notice that in a finite-dimensional case, the iterative sequence generated by the *KA* is always convergent, while the situation in the infinite-dimensional setup may differ. We say that the sequence $(e^n)_{n=0}^{\infty}$ is *effective* if and only if $\forall x \in H \ \ \lim_{n\to\infty} x^n = x$. Kwapień and Mycielski introduced the following sequence $(g^n)_{n=0}^{\infty}$:

$$g^0 := e^0, \text{ and } \forall n \in N,\ g^n := e^n - \sum_{i=0}^{n-1} < e^n, e^i > g^i.$$

Thus, we have $x^n = \sum_{i=0}^{n} < x, g^i > e^i$. They showed that *the sequence* $(e^n)_{n=0}^{\infty}$ *is effective if and only if* $(g^n)_{n=0}^{\infty}$ *is a tight frame with constant 1 for H*. Let us recall that the sequence $(g^n)_{n=0}^{\infty}$ is a *tight frame* with constant 1 if $\| v \|^2 = \sum_{n\in N} \left| < v, g^n > \right|^2$ for each $v \in H$.

In 2005, Haller and Szwarc [17] made a follow up and connected Kwapień-Mycielski results with construction of frames and tight frames in harmonic analysis. They showed that a sequence $(e^n)_{n=0}^{\infty}$ *is an effective sequence if and only if it is linearly dense in H and for a certain matrix* **C** *associated with the Gram matrix of the sequence* $(e^n)_{n=0}^{\infty}$, $\mathbf{C}^*\mathbf{C}$ *is an orthogonal projection, i.e.,* **C** *is partial isometry*.

Recently, Czaja and Tanis pursued further studies [11] concerning the nature of the *KA*. Their starting point was the observation that if a sequence $(e^n)_{n=0}^{\infty}$ is an orthonormal basis in *H*, then $g^n = e^n$ and by the Kwapień-Mycielski theorem, $(e^n)_{n=0}^{\infty}$ is an effective sequence. They introduced the concept of an *almost effective sequence* and characterized such sequences (under the assumption of being a Bessel sequence) in terms of frames. At the beginning of this section, we mentioned that McCormick extended the classical *KA* to solve $\mathbf{A}x = b$, where **A** is a bounded linear transformation on $l^2(N)$ and $b \in \mathrm{Ran}(A)$. McCormick considered a finite-dimensional approximation of such a problem by a sequence of increasing finite-dimensional subspaces. After imposing certain frame conditions on rows of the matrix operator **A**, one can get a better convergence estimate of the process. Thus, assume that an infinite dimensional matrix **A** has rows $a^i \in l^2(N)$ that form a linearly dense system in $l^2(N)$ and choose the initial guess $x^0 := b_0 \dfrac{a^0}{\| a^0 \|^2}$. Czaja and Tanis showed the following: *Let* $\mathbf{A} : l^2(N) \to l^2(N)$ *be a bounded (matrix) transformation. Then, for the initial guess* x^0*, the KA algorithm always converges to a solution if and only if* **A** *is surjective with rows that form an orthogonal basis for* $l^2(N)$.

This note covers only selected contributors to the development of the Kaczmarz method. Other names of particular note include M. Benzi, Y. Censor, P.L. Combettes, S.D. Flåm, F. Natterer, and C. Popa.

5. Epilogue

Professor W. Orlicz once mentioned: “It seems that Cracovian Calculus of Tadeusz Banachiewicz and Kaczmarz method are the most important Polish achievements in numerical analysis between the wars”. In the era of modern computing, *Cracovian Calculus* became obsolete and serves merely as a historical artifact and provides an important example of a non-associative algebra, while the Kaczmarz method is an efficient technique with many prosperous years ahead.

I thank Dr. Danuta Ciesielska from Pedagogical University of Cracow for her help in supplying a copy of the bibliographical item [8]. My thanks also go to Professor Wojtek Czaja from the Norbert Wiener Center, University of Maryland, College Park, USA, for turning my attention to his recent work [11] and discussions on the Kaczmarz method. I am also grateful to an anonymous referee for comments, which improved precision of this paper.

Dedication: I dedicate this paper to my wife Joanna.

References

[1] Bauschke H.H., Borwein J.M., *On projection algorithms for solving convex feasibility problems*, SIAM Rev., 38(3), 1996, 367-426.
[2] Bodewig E., *Bericht über die verschiedenen Methoden zur Lösung eines System linearer Gleichungen mit reellen Koeffizienten. IV, V*, Nederl. Akad. Wetensch., Proc. 51, 1948, 53-64, 211-219.
[3] Briskman J., *Block Kaczmarz Method with Inequalities*, CMC Senior Theses. Paper 862, 2014, http://scholarship.claremont.edu/cmc_theses/862.
[4] Brooks M.A., *A Survey of Algebraic Algorithms in Computerized Tomography*, MSc thesis, University of Ontario Insitute of Technology, 2010.
[5] Byrne C.L., *Iterative Algorithms in Tomography*, University of Massachusetts Lowell, October 2005.
[6] Byrne C.L., *Applied Iterative Methods*, A.K. Peters, Wellesley, MA, 2008.
[7] Cegielski A., *O algorytmie Kaczmarza*, Wiad. Matematyczne, 46 (1), 2010, 27-35.
[8] Cegielski A., *Bibliography on the Kaczmarz method* (up to 2010), www.wmie.uz.zgora.pl/pracownicy/ACegielski/index.html.
[9] Cegielski A., *Iterative Methods for Fixed Point Problems in Hilbert Spaces*, LNM 2057, Springer, 2012.
[10] Cimmino G., *Calcolo approssimato per le soluzioni dei sistemi di equazioni lineari*, La Ricerca Scientifica, XVI, Series II, Anno IX 1, 1938, 326-333.
[11] Czaja W., Tanis J., *Kaczmarz Algorithm and Frames*, International Journal of Wavelets, Multiresolution and Information Processing, September 11(5), 2013, 13 pages.
[12] Dupont M., *Tomographie spectrale à comptage de photons: développement du prototype PIXSCAN et preuve de concept*, Thèse de doctorat, Aix-Marseille Université, Faculté des Sciences, CPPM-T, 2014.
[13] Eldar Y.C., Needell D., *Acceleration of Randomized Kaczmarz Method via the Johnson--Lindenstrauss Lemma*, Numerical Algorithms, 58(2) 2011, 163-177.

[14] Forsythe G.E., *Solving linear algebraic equations can be interesting*, Bull. Amer. Math. Soc. 59, 1953, 299-329.

[15] Gordon R., Bender R., Herman G.T., *Algebraic Reconstruction Techniques (ART) for three--dimensional electron microscopy and X-ray photography*, Journal of Theoretical Biology 29, 1970, 471-481.

[16] Grangeat P., *Analyse d'un systeme d'imagerie 3D par reconstruction à partir de radiographies X en géométrie conique*, These de Doctorat, l'Ecole Nationale Superieure dés Télécommunications, 1987.

[17] Haller R., Szwarc R., *Kaczmarz algorithm in Hilbert paces*, Studia Mathematica, 169(2), 2005, 123-132.

[18] Jamil N., Needell D., Müller J., Lutteroth Ch., Weber G., *Kaczmarz algorithm with soft constraints for user interface layout*, arXiv:1309.7001v2, 2013.

[19] Johnson W.B., Lindenstrauss J., *Extensions of Lipschitz mappings into a Hilbert space*, Conf. in Modern Analysis and Probability, 1984, 189-206.

[20] Kaczmarz S., Steinhaus H., *Theorie der Orthogonalreichen*, Monografie Matematyczne, Tom VI, Warszawa–Lwów 1935.

[21] Kaczmarz S., *Angenäherte Auflösung von Systemen linearer Gleichingen*, Bull. Intern. Acad. Polonaise Sci. Lett., Cl. Sci. Math. Nat. A, 35, 1937, 355-357.

[22] Kaczmarz S., *Approximate solution of systems of linear equations*, International Journal of Control 57, 1993, 1269-1271.

[23] Kiwiel K., *Block-iterative surrogate projection methods for convex feasibility problems*, Linear Algebra Appl., 215, 1995, 225-259.

[24] Kwapień S., Mycielski J., *On the Kaczmarz algorithm of approximation in infinite-dimensional spaces*, Studia Mathematica, 148(1), 2001, 75-86.

[25] Maligranda L., *Stefan Kaczmarz (1895–1939)*, Antiquitates Mathematicae, 1, 2007, 15-62.

[26] McCormick S.F., *The methods of Kaczmarz and row orthogonalization for solving linear equations and least squares problems in Hilbert space*, Indiana Univ. Math. J., 26(6), 1977, 1137-1150.

[27] Needell D., *Randomized Kaczmarz solver for noisy linear systems*, BIT Numerical Mathematics, 50(2), 2010, 395-403.

[28] Needell D., Tropp J.A., *Paved with good intentions: Analysis of randomized block Kaczmarz method*, Linear Algebra and its Applications, 441, 2014, 199-221.

[29] Needell D., Srebro N., Ward R., *Stochastic Gradient Descent, Weighted Sampling, and the Randomized Kaczmarz algorithm*, Math. Progrogramming Series A, to appear.

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

[31] Tanabe K., *Projection method for solving a singular system of linear equations and its applications*, Numer. Math. 17, 1971, 203-214.

[32] Tompkins C., *Projection methods in calculation of some linear problems*, Bull. Amer. Math. Soc. 55, 1949, 520.

[33] Zahavi A., *Super Resoluton project*, ISL Center for Intelligent System, 2001, http://www.cs.technion.ac.il/cis/Project/Projects_done/superResolution/Kaczmarz_real_good_example.htm.