Kaczmarz Projection Algorithms with Rank Two Gain Update

(article starts on next page)

Kaczmarz Projection Algorithms with Rank Two Gain Update

Alexander Stotsky[1]

Abstract

New Kaczmarz algorithms with rank two gain update, extended orthogonality property and forgetting mechanism which includes both exponential and instantaneous forgetting (implemented via a proper choice of the forgetting factor and the window size) are introduced and systematically associated in this paper with well-known Kaczmarz algorithms with rank one update. The parameter convergence was proved using Lyapunov method and convergence of the inverse of the information matrix can be used for further performance improvement. The performance of new algorithms is examined in the problem of estimation of the grid events in the presence of significant harmonic emissions.

Keywords Kaczmarz algorithms in weighted moving window · Exponential & instantaneous forgetting · Updating & downdating · Rank two update versus rank one update · Estimation of the inverse of the information matrix · Extended orthogonality property · Lyapunov method for convergence analysis · Wave form distortion monitoring in smart grids

1 Previous & Related Work

Kaczmarz projection method, [1, 2] is robust and computationally efficient algorithm (alternative to the least squares approach) for estimation of the frequency contents of the oscillating signals in a number of signal processing and system identification applications. In the Kaczmarz algorithm $\theta_k = \theta_{k-1} - \frac{\varphi_k}{\varphi_k^T \varphi_k} [\varphi_k^T \theta_{k-1} - y_k]$ the output of the model $\varphi_k^T \theta_k$ matches exactly the measured signal $y_k = \varphi_k^T \theta_k$ in each discrete step k, which can be obtained by multiplying the parameter vector θ_k by the regressor φ_k^T. The algorithm has the remarkable orthogonality property $\varphi_k^T \tilde{\theta}_k = 0$, where the vector of the parameter mismatch $\tilde{\theta}_k = \theta_k - \theta_*$ (difference between adjustable and true parameters θ_*) is orthogonal to the regressor vector. The square of the length of the harmonic regressor (which consists of the trigonometric functions at different frequencies), $\varphi_k^T \varphi_k$ is constant and depends on the number of frequencies only. The Kaczmarz algorithms were successfully tested in different estimation problems in automotive applications, [3].

Aiming for convergence rate improvement of the Kaczmarz projection method the gain update algorithm was introduced in [4]. Further development was associated with introduction of the forgetting factor which has significant impact on the estimation performance, discounts exponentially old measurements in the gain update and creates a virtual moving window, [5].

The choice of the forgetting factor is associated with the trade-off between rapidity and accuracy of estimation. The gain matrix with forgetting factor is associated in this paper with the window which is moving in time for further improvements. Introduction of the forgetting factor in the sliding window, [6, 7] creates extended forgetting mechanism that includes both exponential and instantaneous forgetting and provides new opportunities for achievement of the trade-off between rapidity and accuracy.

The movement of the window is associated with data updating and downdating that results in recursive updates of the information matrix, which can occur sequentially, [3, 8] or simultaneously, [9]. Sequential updating and downdating results in computationally complex algorithms. The gain update in known Kaczmarz algorithms, [4, 5] is associated with recursive rank one update, (similar to recursive least squares algorithms, [10, 11]) whereas simultaneous updat-

✉ Alexander Stotsky
alexander.stotsky@chalmers.se; alexander.stotsky@telia.com

[1] Systems and Control Division, Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, SE 412 96, Sweden

Table 1 Comparison of Kaczmarz Algorithms with Rank Two and Rank One Updates

Quantity	Kaczmarz Algorithm with Rank Two Update	Kaczmarz Algorithm with Rank One Update
Gain Update	$\Gamma_k = \frac{1}{\lambda} [\Gamma_{k-1} - \Gamma_{k-1} Q_k S^{-1} Q_k^T \Gamma_{k-1}]$	$\Gamma_k = \frac{1}{\lambda} [\Gamma_{k-1} - \frac{\Gamma_{k-1} \varphi_k \varphi_k^T \Gamma_{k-1}}{\lambda + \varphi_k^T \Gamma_{k-1}\varphi_k}]$
Law	$S = \lambda D + Q_k^T \Gamma_{k-1} Q_k$	
Parameter Update	$\theta_k = \theta_{k-1}$	$\theta_k = \theta_{k-1} - \frac{\Gamma_{k-1} \varphi_k}{\varphi_k^T \Gamma_{k-1}\varphi_k} [\varphi_k^T \theta_{k-1} - y_k]$
Law	$- \Gamma_{k-1} Q_k [Q_k^T \Gamma_{k-1} Q_k]^{-1} [Q_k^T \theta_{k-1} - \tilde{y}_k]$	
Inversion Error	$E_k = (I - \Gamma_{k-1} Q_k S^{-1} Q_k^T) E_{k-1}$	$E_k = (I - \frac{\Gamma_{k-1} \varphi_k \varphi_k^T}{\lambda + \varphi_k^T \Gamma_{k-1}\varphi_k}) E_{k-1}$
$E_k = I - \Gamma_k A_k$		
Parameter Error	$\tilde{\theta}_k = (I - \Gamma_{k-1} Q_k [Q_k^T \Gamma_{k-1} Q_k]^{-1} Q_k^T) \tilde{\theta}_{k-1}$	$\tilde{\theta}_k = (I - \frac{\Gamma_{k-1} \varphi_k \varphi_k^T}{\varphi_k^T \Gamma_{k-1}\varphi_k}) \tilde{\theta}_{k-1}$
$\tilde{\theta}_k = \theta_k - \theta_*$		
Orthogonality Property	$Q_k^T \tilde{\theta}_k = 0$	$\varphi_k^T \tilde{\theta}_k = 0$

ing/downdating is associated with computationally efficient rank two update, [12].

New gain update which is explicitly associated with the movement of the sliding window of the size w with exponential forgetting is introduced in Kaczmarz algorithms. The parameter update law extends the property of orthogonality where the output of the model matches measured data in two end points of the weighted window. The development is performed and systematically associated with well-known Kaczmarz algorithms with rank one update, see Table 1 and [5].

It is shown that the gain update Γ_k converges to the inverse of the information matrix and the parameters converge to their true values, which is a new property of Kaczmarz algorithms with rank two update discovered in this paper, see Section 4.

The response time of the estimation algorithms is restricted by the choice of the window size. Short window (which is associated with fast response) implies ill-conditioning of the information matrix which results in sensitivity to numerical calculations and error accumulation. New convergence properties discovered in this paper allow initialization to the approximate inverse with subsequent convergence of the inverse of information matrix, see Section 4. This property opens new opportunities for performance improvement in the ill-conditioned case where the difficulties are associated with the matrix inversion. In addition, Newton-Schulz and Richardson algorithms can be applied for improvements of the transient performance, [12, 13].

The performance of new algorithms is examined in the problem of estimation of the grid events in the presence of significant harmonic emissions, see Fig. 3 and [12–14].

The paper is organized as follows. New Kaczmarz algorithms with rank two update are introduced in the next Section 2. New algorithms are systematically associated with algorithms with rank one update in Section 3 and Table 1. Convergence properties of new algorithms, which is the main contribution of this paper are presented in Section 4. Simulation results which show the advantages of the proposed algorithms are presented in Section 5 and finally Section 6 outlines brief conclusions and directions for further research.

2 Moving Window with Exponential Forgetting

Suppose that the measured oscillating signal is presented in the following form $y_k = \varphi_k^T \theta_*$, where φ_k is the harmonic regressor, $\varphi_k^T = [cos(q_0 k) \; sin(q_0 k) \; ... \; cos(q_h k) \; sin(q_h k)]$, $q_0, ... q_h$ are the frequencies and θ_* is the vector of unknown parameters, $k = 1, 2, ...$

The window of the size w which is moving in time can be presented as rank two update of the information matrix (which is defined as the sum of the outer products of the regressor vectors) where new observation is added (updating) and old observation is deleted (downdating). The estimation performance is highly influenced by the forgetting factor, $0 < \lambda \leq 1$ which discounts exponentially old measurements and creates a virtual window inside of the moving window. In other words, the new data φ_k (with the largest forgetting factor which is equal to one) enter the window and the data with the lowest priority $\tilde{\varphi}_{k-w} = \sqrt{\lambda^w} \varphi_{k-w}$ leave the window in step k as follows, [12]:

$$A_k = \sum_{j=k-(w-1)}^{j=k} \lambda^{k-j} \varphi_j \varphi_j^T$$
$$= \lambda A_{k-1} + Q_k D Q_k^T \tag{1}$$

where $Q_k = [\varphi_k \; \tilde{\varphi}_{k-w}]$, $D = diag[1, -1]$, $k \geq w + 1$.

The gain matrix (as inverse of A_k) which is derived by application of the matrix inversion lemma, [15] to the identity (1) and the parameter update law can be written in the following form, [12]:

$$\Gamma_k = \frac{1}{\lambda} \, [\; \Gamma_{k-1} - \Gamma_{k-1} \; Q_k \; S^{-1} \; Q_k^T \; \Gamma_{k-1} \;] \tag{2}$$

$$\begin{aligned} \theta_k = \theta_{k-1} \\ - \Gamma_{k-1} \; Q_k \; [Q_k^T \; \Gamma_{k-1} \; Q_k]^{-1} \; [Q_k^T \; \theta_{k-1} - \tilde{y}_k] \end{aligned} \tag{3}$$

where $S = \lambda \, D + Q_k^T \; \Gamma_{k-1} \; Q_k$, $\tilde{y}_k^T = [y_k \; \sqrt{\lambda^w} y_{k-w}]$ is the synthetic output, provided that the matrix $Q_k^T \; \Gamma_{k-1} \; Q_k$ is invertible. The vector of adjustable parameters θ_k in (3) is constructed so that the output of the model matches exactly the measured signal $Q_k^T \; \theta_k = \tilde{y}_k$ (see Sections 1 and 3) and the following extended orthogonality property $Q_k^T \; \tilde{\theta}_k = 0$ is valid for the parameter mismatch $\tilde{\theta}_k = \theta_k - \theta_*$.

3 Kaczmarz Algorithms with Rank One Update

Introduction of the forgetting factor allows to establish relationship between Kaczmarz algorithms with rank two and rank one updates. Notice that Q_k gets the following form $Q_k = [\varphi_k \; 0]$ without downdating when $\lambda^w \to 0$. Straightforward substitution of Q_k defined above in (2) yields to following well-known Kaczmarz algorithms, [4, 5]:

$$\Gamma_k = \frac{1}{\lambda} \, [\; \Gamma_{k-1} - \frac{\Gamma_{k-1} \; \varphi_k \; \varphi_k^T \; \Gamma_{k-1}}{\lambda + \varphi_k^T \; \Gamma_{k-1} \varphi_k} \;] \tag{4}$$

$$\theta_k = \theta_{k-1} - \frac{\Gamma_{k-1} \; \varphi_k}{\varphi_k^T \; \Gamma_{k-1} \varphi_k} \; [\varphi_k^T \theta_{k-1} - y_k] \tag{5}$$

with the orthogonality property $\varphi_k^T (\theta_k - \theta_*) = 0$. Comparison of the algorithms (2), (3) and (4), (5) and the error models is presented in Table 1.

The performance of estimation can be further improved using remarkable convergence properties of the algorithms described in the next Section.

4 Convergence of Matrix Inversion and Parameter Errors

Error Models For system (1)–(3) the following error models are valid:

$$E_k = (I - \Gamma_{k-1} \; Q_k \; S^{-1} \; Q_k^T) \; E_{k-1} \tag{6}$$

$$\tilde{\theta}_k = (I - \Gamma_{k-1} \; Q_k \; [Q_k^T \; \Gamma_{k-1} \; Q_k]^{-1} \; Q_k^T) \; \tilde{\theta}_{k-1} \tag{7}$$

where $E_k = I - \Gamma_k A_k$ and $\tilde{\theta}_k = \theta_k - \theta_*$ are matrix inversion and parameter estimation errors (the same matrix inversion error is used in Newton-Schulz algorithms, [13]) and I is the identity matrix.

Convergence of Matrix Inversion The matrices $I - \Gamma_{k-1} \; Q_k \; S^{-1} \; Q_k^T$ can be associated with generalized Householder matrices and explicit solution of E_k can be associated with aggregated Householder transformation (products of Householder matrices). Therefore the convergence of the matrix inversion error for system with harmonic regressor can be established using the arguments similar to [16–18].

Notice that Q_k gets the following form $Q_k = [\varphi_k \; 0]$ for sufficiently small forgetting factor and large window size, i.e. if $\lambda^w \to 0$. Kaczmarz algorithm with rank two update becomes Kaczmarz algorithm with rank one update and the inversion error, (6) gets the form: $E_k = (I - \frac{\Gamma_{k-1} \varphi_k \varphi_k^T}{\lambda + \varphi_k^T \Gamma_{k-1} \varphi_k}) E_{k-1}$, see Table 1. Multiplying this identity by $\varphi_k \varphi_k^T$ and taking the sum (under the assumption that $E_k \approx E_{k-1}$) yields $\sum_{k=r}^{r+w} \varphi_k \varphi_k^T \; E_k = \sum_{k=r}^{r+w} \frac{\lambda}{\lambda + \varphi_k^T \Gamma_{k-1} \varphi_k} \; \varphi_k \; \varphi_k^T \; E_{k-1}$. Thus the convergence of the matrix inversion error can be established for systems with harmonic regressor, [19].

Comparison of the matrix inversion and parameter convergence rates as a function of the forgetting factors for rank one and rank two updates is presented in Fig. 1. Figure 1(a) shows that Kaczmarz algorithm with rank one update provides faster convergence of the inversion error for larger forgetting factors and Fig. 1(b) shows faster convergence of the parameter error for Kaczmarz algorithm with rank two update.

Parameter Convergence The convergence of the parameter vector can be proved using Lyapunov function candidate $V_k = \tilde{\theta}_k^T \; A_k \; \tilde{\theta}_k$ for symmetric and positive definite matrix A_k. Taking into account orthogonality property $Q_k^T \; \tilde{\theta}_k = 0$ evaluation of the first difference $V_k - V_{k-1}$ yields:

$$\begin{aligned} V_k - V_{k-1} &= \lambda \, \tilde{\theta}_k^T \; A_{k-1} \; \tilde{\theta}_{k-1} - V_{k-1} \\ &\quad + \lambda \, \tilde{\theta}_k^T (I - A_{k-1} \Gamma_{k-1}) Q_k [Q_k^T \Gamma_{k-1} Q_k]^{-1} Q_k^T \tilde{\theta}_{k-1} \\ &= -\lambda \, \tilde{\theta}_{k-1}^T \; Q_k \; [Q_k^T \; \Gamma_{k-1} \; Q_k]^{-1} \; Q_k^T \; \tilde{\theta}_{k-1} \\ &\quad - (1 - \lambda) V_{k-1} + \lambda \, (\tilde{\theta}_k + \tilde{\theta}_{k-1})^T \; (I - A_{k-1} \Gamma_{k-1}) \\ &\quad Q_k \; [Q_k^T \; \Gamma_{k-1} \; Q_k]^{-1} \; Q_k^T \; \tilde{\theta}_{k-1} \end{aligned} \tag{8}$$

Notice that the matrix inversion error $(I - A_{k-1}\Gamma_{k-1})$ is vanishing, see Fig. 1(a) and there exist positive constants c and λ_*, $\rho < 1$ such that the following holds $V_k \leq \lambda_* V_{k-1} + c \, \rho^k$ and the bounds

(a)

(b)
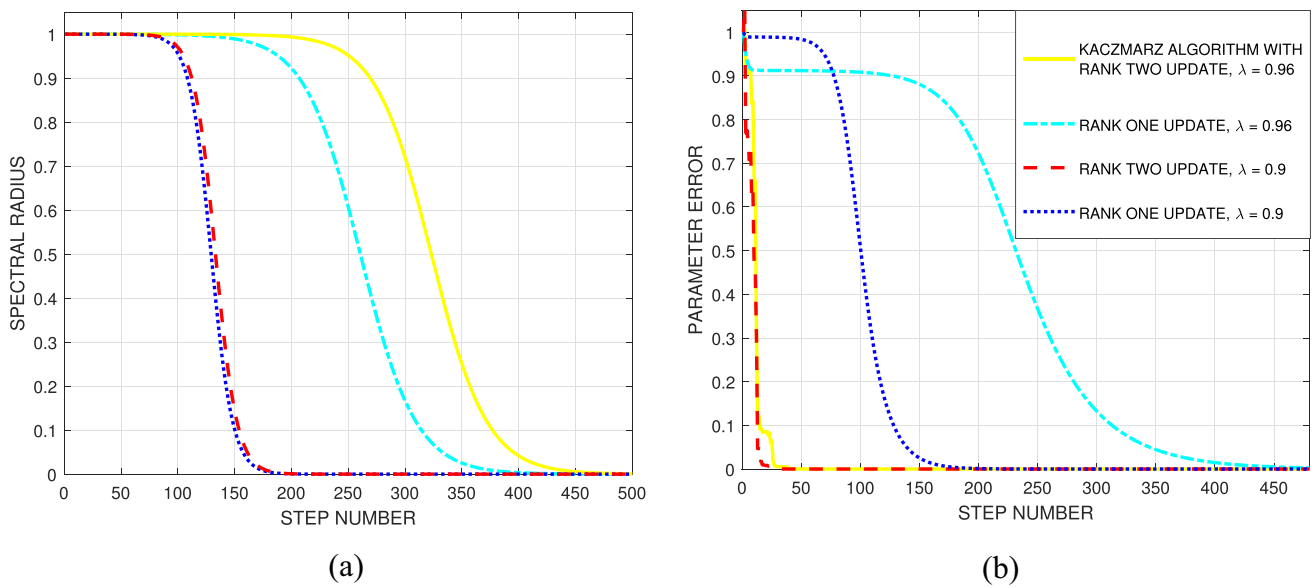
Figure 1 The spectral radius of the matrix inversion error $E_k = I - \Gamma_k A_k$ is plotted in (**a**) for Kaczmarz algorithm with rank two update and window size $w = 30$ with solid yellow line for $\lambda = 0.96$ and red dashed line for $\lambda = 0.9$. The spectral radius for Kaczmarz algorithm with rank one update is plotted with dashdot cyan line for $\lambda = 0.96$ and with dotted blue line for $\lambda = 0.9$. The norms of the parameter errors, $\|\theta_k - \theta_*\|$ are plotted with the same lines and forgetting factors in (**b**).

Figure 2 Rank accumulation for information matrices with harmonic regressor (full rank of information matrix is equal to 16), where $\sum\limits_{k=1}^{k=19} Q_k Q_k^T$ and $\sum\limits_{k=1}^{k=19} \varphi_k \varphi_k^T$ are associated with Kaczmarz algorithms with rank two and rank one updates (red dashed and dotted blue lines respectively).

(a)

(b)
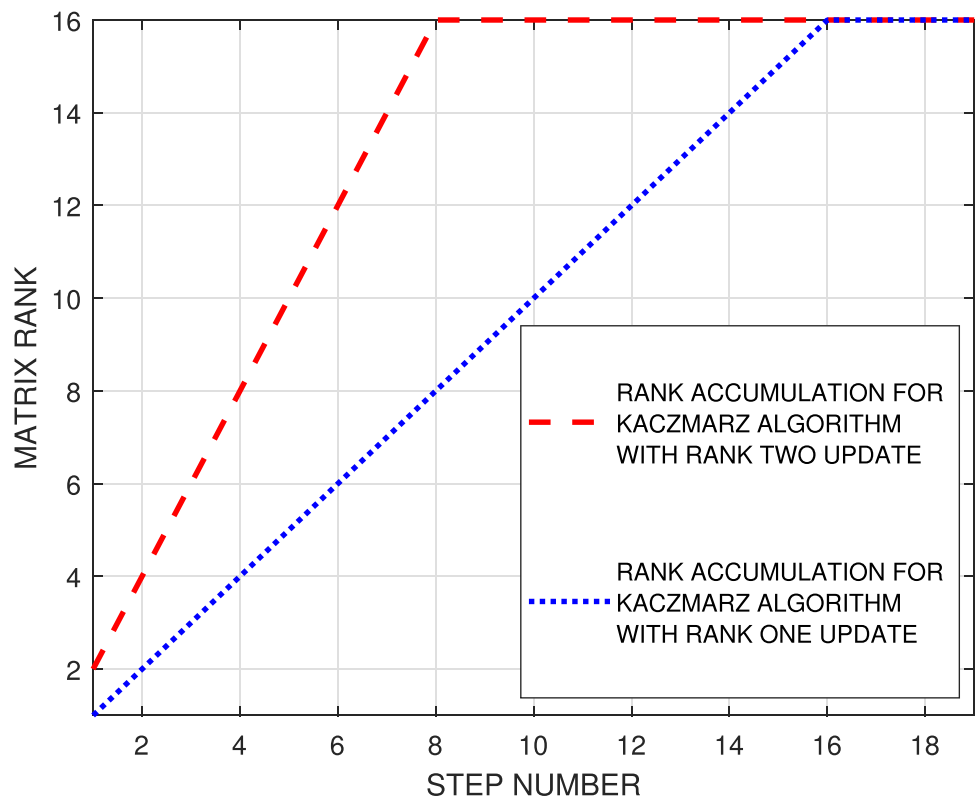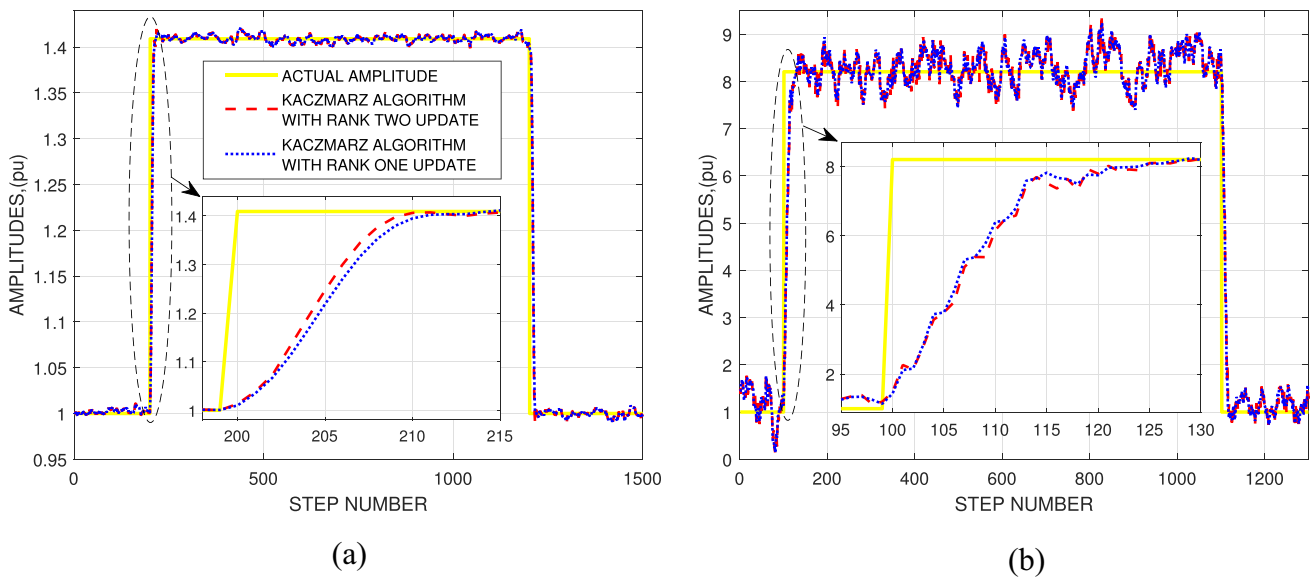
Figure 3 The amplitudes estimated with Kaczmarz algorithms with rank two (2), (3) and rank one (4), (5) updates are plotted with red dashed and dotted blue lines respectively for small window size, $w = 18$ and forgetting factor $\lambda = 0.55$ in the presence of the measurement noise.

Actual amplitudes are plotted with the yellow line. Estimated amplitudes at the frequencies 50 Hz and 150 Hz are plotted in (**a**) and (**b**) respectively.

$$V_k \leq \lambda_*^k \, V_0 + c \, \rho \, \frac{\lambda_*^k - \rho^k}{\lambda_* - \rho} \tag{9}$$

$$V_k \leq \lambda_*^k \, V_0 + c \, k \, \lambda_*^k \tag{10}$$

are valid for $\lambda_* \neq \rho$ and $\lambda_* = \rho$ respectively. Finally, the following bound is true for the parameter error:

$$\|\tilde{\theta}_k\| \leq \sqrt{\frac{V_k}{\lambda_{min}(A_k)}}.$$

Orthogonality property $\tilde{\theta}_k^T \, Q_k = 0$ can also be directly used to establish the parameter convergence. Multiplying this identity by Q_k^T and taking the sum yields $\sum_{k=r}^{r+w} \tilde{\theta}_k^T \, Q_k \, Q_k^T = 0$. Assuming that the vector of adjustable parameters is constant (after some transient) the parameter convergence can be established for persistently exciting regressor. The summation above is associated with rank accumulation using rank two increment. Figure 2 shows that Kaczmarz algorithm with rank one update accumulates the rank slower. Using arguments similar to [19] it can also be shown that the estimates are unbiased in the presence of the noise.

5 Simulations

New algorithms were tested for estimation of the swell signature in the electricity network. The swell event is associated with momentary increase of the voltage that occurs usually

when a large load turns off in a power system. The voltage signal which contains the fundamental frequency of 50 Hz and seven higher harmonics was simulated. The swell events are modeled as momentary increase of the amplitude of the first and the third harmonics, see the yellow lines in Fig. 3. The transient response of the Kaczmarz algorithm with rank two update (2), (3) is compared to the response of the Kaczmarz algorithm with rank one update (4), (5) in the presence of the measurement noise for the same forgetting factor $\lambda = 0.55$ in the short sliding window. The Figure shows that Kaczmarz algorithm with rank two update provides faster estimates at the fundamental frequency for the same forgetting factor.

6 Conclusion

The gain matrix with rank two update which allows exponential weighting of the data inside of the moving window, prioritizes recent measurements and improves estimation performance for fast varying changes of the signal formed the basis for new Kaczmarz algorithms. New Kaczmarz adjustment law has extended orthogonality property in the sliding window and showed improvement of estimation performance in the problem of detection of the grid events. The enhancement of the performance was achieved via a proper choice of two adjustable parameters (small window size and forgetting factor) associated with fast forgetting.

New Kaczmarz algorithms with rank two update were systematically associated in this paper with well-known

Kaczmarz algorithms with rank one update, see Table 1. The parameter convergence was proved using Lyapunov method and the convergence of the inverse of information matrix can be used for further improvement of the performance.

The results are valid under the assumption associated with invertibility of the matrix $Q_k^T \Gamma_{k-1} Q_k$ which can be seen as drawback of proposed algorithm. For the sake of robustness the Kaczmarz algorithms with rank two update can be modified. The modification can be associated with modification of well-known Kaczmarz algorithm, [1].

Declarations

This research was not supported by any organization. This paper does not have any conflict of interest. Participation of the author in this research is completely voluntary. The author gives Springer the permission to publish this paper. This paper does not include any measurement data.

References

1. Åstrom, K. J., & Wittenmark, B. (1989). *Adaptive control*. Addison-Wesley, Reading, Massachusetts.
2. Cegielski, A. (2024). Bibliography on the Kaczmarz method. Institute of Mathematics, University of Zielona Gora, Zielona Gora, Poland. Retrieved January 8, 2024, from https://staff.uz.zgora.pl/acegiels/Publications-Kaczmarz-method.pdf
3. Stotsky, A. (2009). *Automotive engines: Control estimation, statistical detection*. Berlin-Heidelberg: Springer-Verlag.
4. Avedyan, E. D. (1978). Modified Kaczmarz algorithms for estimating the parameters of linear plants. *Automation and Remote Control, 39*(5), 674–680.
5. Stotsky, A. (2012). A new frequency domain system identification method. *Proceedings of the IMechE Part I: Journal of Systems and Control Engineering, 226*(1), 111–124.
6. Liu, H., & He, Z. (1995). A sliding-exponential window RLS adaptive filtering algorithm: Properties and applications. *Signal Processing, 45*, 357–368.
7. Djigan, V. (2006). Multichannel parallelizable sliding window RLS and fast RLS algorithms with linear constraints. *Signal Processing, 86*, 776–791.
8. Zhang, Q. (2000). Some implementation aspects of sliding window least squares algorithms. *IFAC Proceedings Volumes, 33*(15), 763–768. https://www.sciencedirect.com/science/article/pii/S1474667017398440?via%3Dihub
9. Choi, J., & Boyanczyk, A. (1995). Recursive least squares problems on distributed memory mupliprocessors. *Journal of Parallel and Distributed Computing, 24*, 11–26.
10. Fomin, V., Fradkov, A., & Yakubovich, V. (1981). *Adaptive control of dynamic plants*. Moscow: Nauka. In Russian.
11. Ljung, L., & Söderström, T. (1983). *Theory and practice of recursive identification* The MIT Press Series in Signal Processing, Optimization, and Control (Vol. 4). MIT Press.
12. Stotsky, A. (2023). Recursive estimation in the moving window: Efficient detection of the distortions in the grids with desired accuracy. *Journal of Advances in Applied and Computational Mathematics, 9*, 181–191.
13. Stotsky, A. (2022). Recursive versus nonrecursive Richardson algorithms: Systematic overview, unified frameworks and application to electric grid power quality monitoring. *Automatika, 63*(2), 328–337.
14. Stotsky, A. (2022). Simultaneous frequency and amplitude estimation for grid quality monitoring: New partitioning with memory based Newton-Schulz corrections. *IFAC PapersOnLine, 55*(9), 42–47.
15. Hager, W. (1989). Updating the inverse of a matrix. *SIAM Review, 31*(2), 221–239.
16. Rader C., & Steinhard A. (1986, December). Hyperbolic householder transformations. *IEEE Transactions on Acoustics, Speech, and Signal Processing, ASSP-34*(6), 1589–1602.
17. Ljung, L., & Gunnarsson, S. (1990). Adaptive tracking in system identification - a survey. *Automatica, 26*(1), 7–22.
18. Guo, L. (1994). Stability of recursive stochastic tracking algorithms. *SIAM Journal on Control and Optimization, 32*, 1195–1225.
19. Stotsky, A. (2010). Recursive trigonometric interpolation algorithms. *Journal of Systems and Control Engineering, 224*(1), 65–77.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.