Recursive least squares estimation with rank two updates

Citation for the original published paper (version of record):
Stotsky, A. (2025). Recursive least squares estimation with rank two updates. Automatika, 66(4): 619-624. http://dx.doi.org/10.1080/00051144.2025.2517431

Taylor & Francis
Taylor & Francis Group

Automatika

Journal for Control, Measurement, Electronics, Computing and Communications

ISSN: 0005-1144 (Print) 1848-3380 (Online) Journal homepage: www.tandfonline.com/journals/taut20

Recursive least squares estimation with rank two updates

Alexander Stotsky

AUTOMATIKA
2025, VOL. 66, NO. 4, 619–624
https://doi.org/10.1080/00051144.2025.2517431

Recursive least squares estimation with rank two updates

Alexander Stotsky

Systems and Control, Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden

ABSTRACT

This paper presents new recursive least squares (RLS) algorithms with enhanced performance, achieved via a combination of exponential forgetting and windowing techniques. The proposed algorithms with rank two updates are systematically aligned with established RLS algorithms with rank one updates to ensure unification and clarity. Newly identified properties of the recursive algorithms, associated with the convergence of both the inverse of the information matrix and the parameter estimates which are presented in this paper, offer great potential for further enhancement of the estimation performance. The proposed algorithms demonstrate significant improvements in the estimation of the grid events in the presence of substantial harmonic emissions.

ARTICLE HISTORY
Received 29 January 2024
Accepted 29 May 2025

KEYWORDS
Forgetting and windowing; updating and downdating; rank two update versus rank one update; RLSR2 (recursive least squares with rank two update); estimation of the inverse of the information matrix and unknown parameters via RLS algorithms; wave form distortion monitoring for smart grids

1. Motivation and main results

RLS algorithms with exponential forgetting are broadly utilized in system identification, signal processing, statistics, adaptive control, and in a wide range of other applications [1–3].

A fixed forgetting factor, being the sole adjustable parameter, is insufficient to guarantee satisfactory estimation performance across all scenarios, which has led to numerous modifications of the RLS algorithms, see for example [4,5] and references therein.

Improved performance is achieved via the integration of exponential forgetting and windowing techniques, with the forgetting factor and the window size serving as adjustable parameters which enable optimized trade-offs and facilitate faster response to rapid signal changes, [6–9].

Window movement entails both updating and downdating events, leading to recursive rank two updates of the information matrix. This contrasts with the classical RLS algorithm, which relies solely on recursive rank one updates, [1–3].

This paper extends the sliding window least squares approach outlined in Ref. [10] by integrating exponential forgetting, thereby prioritizing recent observations and improving estimation performance for rapidly changing signals.

The development aligns systematically with standard RLS algorithms with rank one updates, see Table 1. Furthermore, the rank two gain update, derived from the least squares formulation with exponential forgetting in the sliding window, forms the basis of improved Kaczmarz algorithms, [8].

This paper reveals a novel property of RLS algorithms with rank two updates: the gain matrix Γ_k converges to the inverse of the information matrix, and the parameter estimates converge to their true values, see Figure 2 in Section 3. The transient performance of the algorithms can be further enhanced through the application of Newton-Schulz and Richardson corrections [10–12].

The efficiency of the new algorithms is assessed through their application to grid event estimation in the presence of substantial harmonic emissions, see Figure 3 and [10–13].

2. RLS algorithms

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, $\varphi_k^T = [\cos(q_0 k)\ \sin(q_0 k) \cdots \cos(q_h k)\ \sin(q_h k)]$, $q_0, \ldots q_h$ are the frequencies and θ_* is the vector of unknown parameters, $k = 1, 2, \ldots$. The oscillating signal y_k is approximated by the model $\hat{y}_k = \varphi_k^T \theta_k$.

The parameter vector θ_k is determined via the least squares method, which entails minimization of the

CONTACT Alexander Stotsky alexander.stotsky@chalmers.se, alexander.stotsky2@telia.com Systems and Control, Department of Electrical Engineering, Chalmers University of Technology, Gothenburg SE-412 96, Sweden

Table 1. RLS Algorithms with rank two and rank one updates, see also Table 1 in Ref. [8] for comparison with the Kaczmarz algorithms.

Variable	RLS Algorithm with Rank Two Update	RLS Algorithm with Rank One Update
Gain Update	$\Gamma_k = \frac{1}{\lambda}[\Gamma_{k-1} - \Gamma_{k-1}Q_kS^{-1}Q_k^T\Gamma_{k-1}]$	$\Gamma_k = \frac{1}{\lambda}\left[\Gamma_{k-1} - \frac{\Gamma_{k-1}\varphi_k\varphi_k^T\Gamma_{k-1}}{\lambda + \varphi_k^T\Gamma_{k-1}\varphi_k}\right]$
Parameter Update	$\theta_k = \theta_{k-1} - \Gamma_{k-1}Q_kS^{-1}[Q_k^T\theta_{k-1} - \tilde{y}_k]$	$\theta_k = \theta_{k-1} - \frac{\Gamma_{k-1}\varphi_k}{\lambda + \varphi_k^T\Gamma_{k-1}\varphi_k}[\varphi_k^T\theta_{k-1} - y_k]$
Inversion Error	$E_k = (I - \Gamma_{k-1}Q_kS^{-1}Q_k^T)E_{k-1}$	$E_k = \left(I - \frac{\Gamma_{k-1}\varphi_k\varphi_k^T}{\lambda + \varphi_k^T\Gamma_{k-1}\varphi_k}\right)E_{k-1}$
$E_k = I - \Gamma_kA_k$		
Parameter Error	$\tilde{\theta}_k = (I - \Gamma_{k-1}Q_kS^{-1}Q_k^T)\tilde{\theta}_{k-1}$	$\tilde{\theta}_k = \left(I - \frac{\Gamma_{k-1}\varphi_k\varphi_k^T}{\lambda + \varphi_k^T\Gamma_{k-1}\varphi_k}\right)\tilde{\theta}_{k-1}$
$\tilde{\theta}_k = \theta_k - \theta_*$		

following loss function:

$$P_k = \sum_{j=k-(w-1)}^{k} \lambda^{k-j}\,(y_j - \varphi_j^T\theta_k)^2 \tag{1}$$

where the forgetting factor $0 < \lambda < 1$ downweights older data within the window of the size w. All observations are weighted equally, provided that the forgetting factor is equal to one, $\lambda = 1$.

Minimization of the loss function (1) yields the following system of algebraic equations:

$$A_k\theta_k = b_k, \quad A_k = \sum_{j=k-(w-1)}^{j=k} \lambda^{k-j}\varphi_j\,\varphi_j^T \tag{2}$$

$$b_k = \sum_{j=k-(w-1)}^{j=k} \lambda^{k-j}\varphi_j y_j \tag{3}$$

which should be solved with respect to the vector of estimated parameters θ_k in each step k. The information matrix A_k in (2) can also be defined as the rank two update of the matrix A_{k-1}, $k \geq w+1$. Rank two updates are associated with the movement of the window, where a new observation is added (updating) and an old observation is removed (downdating), [10].

In other words, the new data φ_k, y_k (with the largest forgetting factor which is equal to one) enter the window and the data with the lowest priority $\tilde{\varphi}_{k-w} = \sqrt{\lambda^w}\,\varphi_{k-w}$, $\lambda^w\varphi_{k-w}y_{k-w}$ leave the window in step k:

$$A_k = \lambda A_{k-1} + Q_k\,D\,Q_k^T, \quad b_k = \lambda b_{k-1} + d_k \tag{4}$$

where $Q_k = [\varphi_k\;\tilde{\varphi}_{k-w}]$, $D = \text{diag}[1, -1] = \left[\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right]$ and $d_k = \underbrace{\varphi_k y_k}_{\text{updating}} - \underbrace{\lambda^w\varphi_{k-w}y_{k-w}}_{\text{downdating}}$.

Notice that the matrix Q_k contains scaled regressor $\tilde{\varphi}_{k-w}$ in order to avoid singularity in the case where $\lambda^w \to 0$ for sufficiently small λ and sufficiently large w. Inclusion of sufficiently small λ^w in the matrix D (without scaling the regressor) makes this matrix singular, which results in large estimation errors when calculating the inverse of A_k. Notice also that the rank one update can be obtained as the limiting form of rank two update (4) with $\lambda^w \to 0$, see Table 1 and [1–3].

The RLS algorithms with rank two update and exponential forgetting which are derived by application of the matrix inversion lemma to the identity (4) can be written in the following form [8,10]:

$$\Gamma_k = \frac{1}{\lambda}[\Gamma_{k-1} - \Gamma_{k-1}Q_kS^{-1}Q_k^T\Gamma_{k-1}] \tag{5}$$

$$\theta_k = \theta_{k-1} - \Gamma_{k-1}Q_kS^{-1}[Q_k^T\theta_{k-1} - \tilde{y}_k] \tag{6}$$

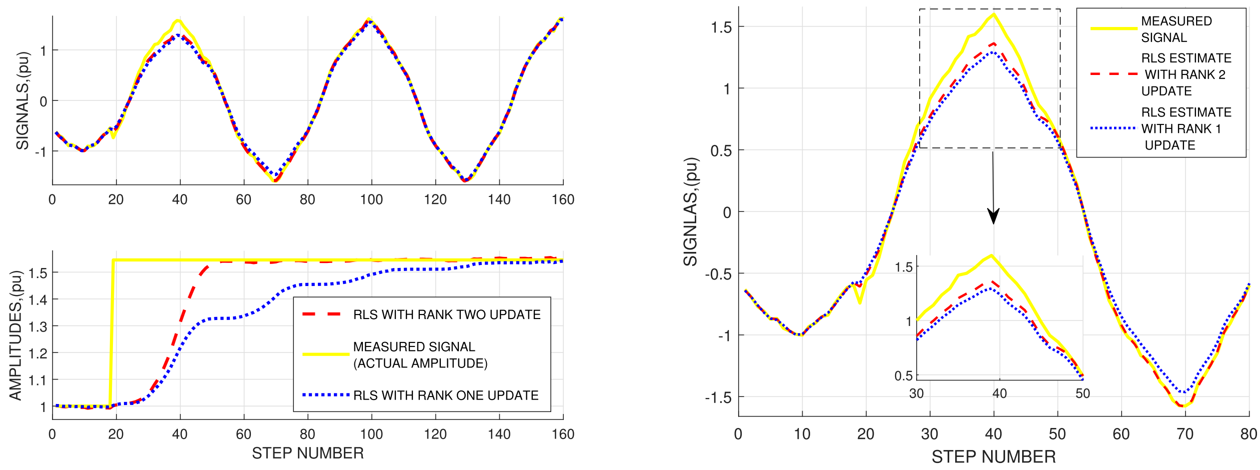
where $S = \lambda D + Q_k^T\Gamma_{k-1}Q_k$ and $\tilde{y}_k^T = [y_k\,\sqrt{\lambda^w}y_{k-w}]$ is the synthetic/extended output, provided that the matrix $Q_k^T\Gamma_{k-1}Q_k$ is invertible. Initialization $\Gamma_w = A_w^{-1}$ and $A_w\theta_w = b_w$ implies that the algorithm (5) and (6) guarantees that the following equations $\Gamma_k = A_k^{-1}$, $\theta_k = A_k^{-1}b_k$ hold in each step k. Notice that the algorithm (5) and (6) (similar to algorithm (7) and (8)) has the convergence properties of the matrix inversion and parameter errors, see Table 1 and Section 3. These properties can be utilized for improvement of the estimation performance.

Interestingly enough that similar form of the estimation algorithms for the vector output sequence and regressor matrix was studied in Refs [14,15]. Notice that the multiple output systems considered in Refs [14,15] are not associated with the sliding window technique, updating/downdating events and extended forgetting mechanism and therefore cannot be applied to the single output case with rank two update.

Notice that the least squares solution in sliding window which resulted in algorithms (5) and (6) formed the basis for new Kaczmarz projection method, which is described in Ref. [8]. RLS algorithms (5) and (6) can also be seen as a robust modification of the Kaczmarz algorithms with rank two gain update.

Notice also that recursive calculations of the coefficients of Discrete Fourier Transform in moving window (which is a special case of RLS algorithms) can be presented in a very simple form, [16].

Finally, this approach can be extended to the continuous-time case by leveraging the concepts introduced in Refs [17,18].



(a) (b)

Figure 1. The estimate of the amplitude of the first harmonic calculated with the algorithm (5) and (6) (red dashed line) is compared in the second plot of Figure 1(a) to the estimate provided by the algorithm (7) and (8) (dotted blue line) with the same forgetting factor $\lambda = 0.97$. The actual change of the amplitude of the first harmonic due to swell event is plotted with the yellow line in the second plot of Figure 1(a). The performance of approximation of the measured signal (the yellow line) by the signals calculated by algorithms (5), (6) and (7), (8) (the red dashed and dotted blue line, respectively) is shown in the first plot of Figure 1(a) and in Figure 1(b).

Relation between algorithms and rank reduction. Introduction of the forgetting factor allows to establish relationship between RLS algorithms with rank two and rank one updates. Notice that Q_k and d_k, see (4), get the following forms $Q_k = [\varphi_k \ 0]$, $d_k = \varphi_k y_k$ and $\tilde{y}_k^T = [y_k \ 0]$ if $\lambda^w \to 0$ which corresponds to cancellation of downdating. Straightforward substitution of Q_k and $\tilde{y}_k$ defined above in (5) and (6) yields to the following well-known RLS algorithms, [1–3]:

$$\Gamma_k = \frac{1}{\lambda}\left[\Gamma_{k-1} - \frac{\Gamma_{k-1}\varphi_k\varphi_k^T\Gamma_{k-1}}{\lambda + \varphi_k^T\Gamma_{k-1}\varphi_k}\right] \tag{7}$$

$$\theta_k = \theta_{k-1} - \frac{\Gamma_{k-1}\varphi_k}{\lambda + \varphi_k^T\Gamma_{k-1}\varphi_k}[\varphi_k^T\theta_k - y_k] \tag{8}$$

The error models for the algorithms (5), (6) and (7), (8) are presented in Table 1.

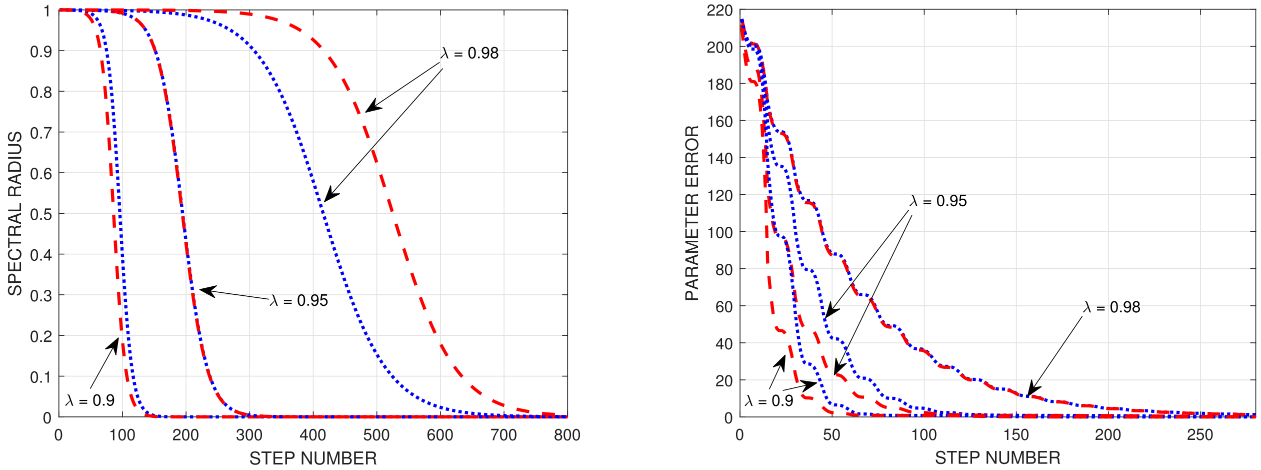
2.1. Simulations

RLS algorithms were tested for detecting voltage swell events in power networks. These events, defined by abrupt voltage increases, can degrade power quality and cause significant financial losses. Effective monitoring of such disturbances is essential in power system engineering.

The swell event is modelled as momentary increase of the amplitude of the first harmonic, see the yellow line in the second plot of Figure 1(a). The transient response of the algorithm (5) and (6) is compared to the response of the algorithm (7) and (8) for the same forgetting factor $\lambda = 0.97$. The amplitudes estimated with algorithms (5), (6) and (7), (8) are plotted in Figure 1(a,b) with the red dashed and dotted blue lines, respectively, whereas the measured signal is plotted with the yellow line.

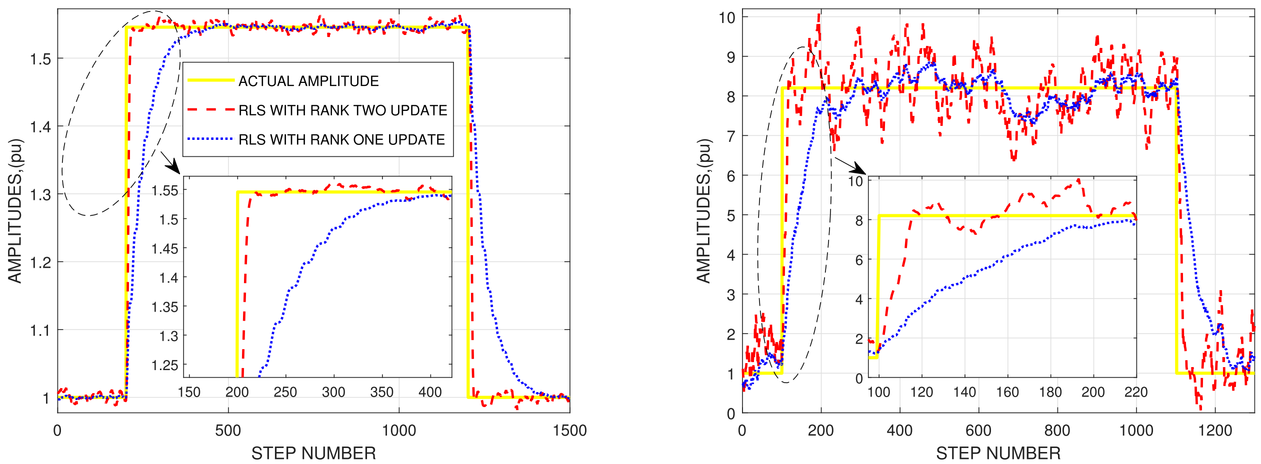
New algorithm (5) and (6) provides faster estimation compared to known RLS algorithm (7) and (8) for the same forgetting factor, see Figure 1(a,b). However, approximately the same fast transient performance can be achieved by reduction of the forgetting factor in known RLS algorithm (7) and (8) or by reduction of the window size in algorithms described in Ref. [10] with forgetting factor which is equal to one.

Notice that both fast forgetting and small window size imply large condition number of the information matrix, sensitivity to measurement noise, numerical operations and significant error accumulation. Algorithm (5) and (6) has two adjustable parameters (the window size w and the forgetting factor λ) which provide additional opportunities for optimization (in comparison to (7) and (8) and algorithms described in Ref. [10]) and hence for performance improvement. The choice of both forgetting factor and the window size is associated with the tradeoff between rapidity of estimation and both accuracy (sensitivity to measurement noise) and the condition number. On the one hand, the forgetting factor and the window size should be small enough for rapid estimation. On the other hand, the window size should be large enough and the forgetting factor should be close to one for reduction of the condition number and error accumulation as well as for attenuation of the measurement noise.



(a) (b)

Figure 2. The spectral radius of the matrix inversion error $E_k = I - \Gamma_k A_k$ is plotted in Figure 2(a) with red dashed and dotted blue lines for RLS algorithm with rank two and rank one updates, respectively. The simulations were performed for the following forgetting factors: $\lambda = 0.98$, $\lambda = 0.95$ and $\lambda = 0.9$. The norms of the parameter errors, $\|\theta_k - \theta_*\|$ are plotted with the same lines and forgetting factors in Figure 2(b).



(a) (b)

Figure 3. The amplitudes estimated with RLS algorithms with rank two and rank one updates are plotted with red dashed and dotted blue lines, respectively, for small window size $w = 18$ and forgetting factor $\lambda = 0.98$. Actual amplitudes are plotted with the yellow line. Estimated amplitudes at the frequencies 50 Hz and 150 Hz are plotted in Figure 3(a,b), respectively.

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

3. Convergence of matrix inversion and parameter errors

Error Models. For system (4)–(6) the following error models are valid:

$$E_k = (I - \Gamma_{k-1} Q_k S^{-1} Q_k^T) E_{k-1} \tag{9}$$

$$\tilde{\theta}_k = (I - \Gamma_{k-1} Q_k S^{-1} Q_k^T) \tilde{\theta}_{k-1} \tag{10}$$

where $\tilde{\theta}_k = \theta_k - \theta_*$ is the parameter estimation error, $E_k = I - \Gamma_k A_k$ is the matrix inversion error, where Γ_k is the estimate of the inverse of A_k, [12] and I is the identity matrix.

Convergence of Matrix Inversion. Explicit solution of the Equation (9) is associated with products of the matrices, $I - \Gamma_{k-1} Q_k S^{-1} Q_k^T$ and the convergence can be established for systems with harmonic regressor using the techniques described for example in Refs [19–22].

In other words, multiplying identity (9) by $Q_k Q_k^T$ and taking the sum (under the assumption that $E_k \approx E_{k-1}$) yields $\sum_{k=r}^{r+w} Q_k Q_k^T E_k = \lambda \sum_{k=r}^{r+w} Q_k \;\; D \;\; S^{-1} Q_k^T E_{k-1}$. The matrices in the both sides of the equation are SDD (Strictly Diagonally Dominant) matrices, [23], whereas the diagonal elements in the left hand side increase faster as a function of the window size. Thus the convergence of the matrix inversion error can be established for systems with harmonic regressor.

Notice that a similar error model, which guarantees the convergence is valid for RLS algorithm with rank one update, see Table 1 and [1–3,8].

Comparison of convergence of matrix inversion and parameter errors as function of the forgetting factors for rank two (red dashed line) and rank one (dotted blue line) updates is presented in Figure 2. Figure 2(a) shows that RLS algorithm with rank one updates provides faster convergence of the inversion error for larger forgetting factors and Figure 2(b) shows faster convergence of the parameter error for RLS algorithm with rank two updates.

Parameter Convergence. Transient parameter estimation performance is evaluated (using arguments similar to Refs [3,8,24–26]) via the first difference $V_k - V_{k-1}$ of the Lyapunov function $V_k = \tilde{\theta}_k^T A_k \tilde{\theta}_k$ as follows:

$$V_k = \lambda V_{k-1} - \lambda \tilde{\theta}_{k-1}^T Q_k S^{-1} Q_k^T \tilde{\theta}_{k-1} + \lambda(\tilde{\theta}_k + \tilde{\theta}_{k-1})^T (I - A_{k-1}\Gamma_{k-1}) Q_k S^{-1} Q_k^T \tilde{\theta}_{k-1} \tag{11}$$

Notice that the matrix inversion error $(I - A_{k-1}\Gamma_{k-1})$ is vanishing, see Figure 2(a) and there exist positive constants c and $\lambda_*, \rho < 1$ such that the following hold:

$$V_k \leq \lambda_* V_{k-1} + c\rho^k \tag{12}$$

$$V_k \leq \lambda_*^k V_0 + c\rho \frac{\lambda_*^k - \rho^k}{\lambda_* - \rho} \tag{13}$$

where the bound (13) is valid for $\lambda_* \neq \rho$, [8]. Finally, the following bound is true for the parameter mismatch: $\|\tilde{\theta}_k\| \leq \sqrt{\frac{V_k}{\lambda_{\min}(A_k)}}$.

Transient estimation performance (in the simulated system with harmonic regressor and significant measurement noise) of RLS algorithms at different frequencies in the short window is shown in Figure 3. RLS algorithm with rank two updates provides faster but more noisy estimates (especially at higher frequencies) for the same forgetting factor.

4. Conclusions and outlook

Integrating exponential forgetting into the sliding window significantly enhanced the tracking performance of RLS algorithms. It was also demonstrated that the assessment of the swell signature in power systems can be significantly enhanced via optimal selection of the window size and the forgetting factor.

The unified description of RLS algorithms presented in Table 1 may serve as a foundation for further systematization of system identification algorithms and the development of algorithms with enhanced performance.

Finally, parameter estimation performance can be improved via convergence rate improvement of the matrix inversion error $I - \Gamma_k A_k$.

Funding

This research was not supported by any organization. Open access funding provided by Chalmers University of Technology.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

[1] Fomin V, Fradkov A, Yakubovich V. Adaptive control of the dynamic plants. Moscow: Nauka; 1981. in Russian.

[2] Ljung L, Söderström T. Theory and practice of recursive identification. Cambridge (MA): MIT Press; 1983. (The MIT press series in signal processing, optimization, and control; Vol. 4).

[3] Åstrom KJ, Wittenmark B. Adaptive control. Reading (MA): Addison-Wesley; 1989.

[4] Salgado M, Goodwin G, Middleton R. Modified least squares algorithm incorporating exponential resetting and forgetting. Int J Control. 1988;47(2):477–491. doi: 10.1080/00207178808906026

[5] Lai B, Bernstein D. Generalized forgetting recursive least squares: stability and robustness guarantees. IEEE Trans Autom Control. 2024;69(11):7646–7661. doi: 10.1109/TAC.2024.3394351

[6] Liu H, He Z. A sliding-exponential window RLS adaptive filtering algorithm: properties and applications. Signal Process. 1995;45:357–368. doi: 10.1016/0165-1684(95)00063-J

[7] Djigan V. Multichannel parallelizable sliding window RLS and fast RLS algorithms with linear constraints. Signal Process. 2006;86:776–791. doi: 10.1016/j.sigpro.2005.06.010

[8] Stotsky A. Kaczmarz projection algorithms with rank two gain update. J Signal Process Syst. 2024;96(4–5):327–332. doi: 10.1007/s11265-024-01915-w

[9] Stotsky A. Detection and control of credit card fraud attacks in sliding window with exponential forgetting. Int J Comput Appl (0975–8887). 2025;186(74):9–15.

[10] Stotsky A. Recursive estimation in the moving window: efficient detection of the distortions in the grids with desired accuracy. J Adv Appl Comput Math. 2023;9:181–191. doi: 10.15377/2409-5761.2022.09.14

[11] Stotsky A. Accuracy improvement in least-squares estimation with harmonic regressor: new preconditioning and correction methods. In: 54-th CDC, Dec. 15–18; Osaka, Japan: 2015. p. 4035–4040.

[12] Stotsky A. Recursive versus nonrecursive richardson algorithms: systematic overview, unified frameworks and application to electric grid power quality monitoring. Automatika. 2022;63(2):328–337. doi: 10.1080/00051144.2022.2039989

[13] Stotsky A. Simultaneous frequency and amplitude estimation for grid quality monitoring: new partitioning with memory based Newton–Schulz corrections. IFAC PapersOnLine. 2022;55(9):42–47. doi: 10.1016/j.ifacol.2022.07.008

[14] Islam S, Bernstein D. Recursive least squares for real-time implementation. IEEE Control Syst Mag. 2019;39(3):82–85. doi: 10.1109/MCS.5488303
[15] Brüggemann S, Bitmead R. Exponential convergence of recursive least squares with forgetting factor for multiple-output systems. Automatica. 2021;124(2): 109389. doi: 10.1016/j.automatica.2020.109389
[16] Stotsky A, Kolmanovsky I. Computationally efficient filtering algorithms for engine torque estimation. In: 2005 American Control Conference, June 8–10; Portland (OR): 2005, pp. 5035–5040.
[17] Stotsky A. Lyapunov design for convergence rate improvement in adaptive control. Int J Control. 1993; 57(2):501–504. doi: 10.1080/00207179308934403
[18] Pan Y, Shi T. Adaptive estimation and control with online data memory: a historical perspective. IEEE Control Syst Lett. 2024;8:267–278. doi: 10.1109/LCSYS.2024.3364588
[19] Bischof C, Van Loan C. The WY representation for products of housholder matrices, TR 85-681, Department of Computer Science, Cornell University, Ithaca (NY), 1985.
[20] Rader C, Steinhard A. Hyperbolic householder transformations. IEEE Trans Acoust Speech Signal Process. 1986;ASSP-34(6):1589–1602. doi: 10.1109/TASSP.1986.1164998
[21] Eirola T, Nevanlinna O. Accelerating with rank-one updates. Linear Algebra Appl. 1989;121:511–520. doi: 10.1016/0024-3795(89)90719-2
[22] Guo L. Stability of recursive stochastic tracking algorithms. SIAM J Control Optim. 1994;32:1195–1225. doi: 10.1137/S0363012992225606
[23] Stotsky A. Recursive trigonometric interpolation algorithms. J Syst Control Eng. 2010;224(1):65–77.
[24] Johnstone R, Johnson C, Bitmead R, et al. Exponential convergence of recursive least squares with exponential forgetting factor. Syst Control Lett. 1982;2(2):77–82. doi: 10.1016/S0167-6911(82)80014-5
[25] Stotsky A. A new frequency domain system identification method. Proc Inst Mech Eng Part I: J Syst Control Eng. 2012;226(1):111–124.
[26] Stotsky A. Harmonic regressor: robust solution to least-squares problem. Proc Inst Mech Eng Part I: J Syst Control Eng. 2013;227(8):662–668.