

Rekurzivni najmanjši kvadrati z domensko usmerjenim pozabljanjem

Gregor Černe, Igor Škrjanc

Fakulteta za elektrotehniko, Univerza v Ljubljani
E-pošta: gregor.cerne@fe.uni-lj.si, igor.skrjanc@fe.uni-lj.si

Recursive least squares with domain-driven variable-direction forgetting

Recursive least squares (RLS) assumes persistent excitation; however, in closed-loop systems this condition is often not met, as control quality takes priority over excitation. We show that the state-of-the-art approach—RLS with variable-direction forgetting (VDF)—is suboptimal in two ways: it may forget the most of the covariance matrix even when the excitation becomes singular, and it introduces an additional hyperparameter that complicates the design of complex systems.

To overcome these shortcomings, we propose domain-driven variable-direction forgetting (RLS-3DF). Instead of minimizing the error over samples for which persistent excitation cannot be assumed, RLS-3DF is derived from the expected squared error (ESE) of an RLS model over the important part of the input domain. Shifting from discrete samples to a continuous formulation decouples the sampling process from the domain on which accuracy is required.

The results are promising: RLS-3DF improves upon RLS-VDF by at least 20 % across all examples, including in persistently excited scenarios.

1 Uvod

Osnovna metodologija pri identifikaciji sistemov temelji na optimizaciji parametrov z metodo najmanjših kvadratov (LS). Ključna prednost metode LS je njena nizka računska kompleksnost, saj je rešitev analitična (v nasprotju z metodami gradientnega spusta). Zaradi te lastnosti je rekurzivna metoda najmanjših kvadratov (RLS) uporabna za učenje v realnem času v dinamičnih okoljih, kjer se mora model hitro prilagajati spremembam. Primeri uporabe RLS vključujejo regresijo nad podatkovnimi tokovi [1], adaptivno vodenje [2], prediktivno vodenje [3], načrtovanje eksperimentov [4], nadzor in zaznavanje napak [5] ter napovedno vzdrževanje [6], med drugim.

Čeprav je metoda zelo vsestranska, je bilo v literaturi o identifikaciji sistemov predlaganih več prilagoditev RLS z namenom odprave njenih pomanjkljivosti, kot so občutljivost na šum in izstopajoče vrednosti [1], šum na vhodih [1], predpostavljen avtoregresivni šum [4], korelirane spremenljivke [4], kompromis med stabilnostjo

in sledenjem [7] ter prilagajanje faktorja pozabljanja z uporabo RLS z variabilnim faktorjem pozabljanja [7].

Čeprav so bile te pomanjkljivosti do določene mere odpravljene, odprava zahteve po vztrajnem vzburjanju ostaja aktivno raziskovalno področje. Natančneje, kadar regresijski vektorji niso dovolj raznoliki, lahko signali postanejo korelirani, razmerje med signalom in šumom pa v določenih smereh vhodne domene premajhno, kar vodi v divergenco parametrov. Težava je bila obravnavana večkrat, z mešanimi rezultati. Pristopi segajo od enostavnega praga na Mahalanobisovi razdalji [8], ponastavitve kovariančne matrike, in vse do state-of-the-art variabilno usmerjenega pozabljanja (VDF) [9], ki pozablja le v smereh, ki so primerno vzburjene. V zelo kratkem času (od leta 2020) je bil uporabljen na mnogih področjih, kot so analiza možganskih valov [10], vodenje helikopterja [11], oceni mase vozila [12] in optimizaciji intenzifikacije inzulina pri sladkorni bolezni [13]. Kljub temu pa večina teh algoritmov temelji na kriterijski funkciji, ki je definirana kot vsota kvadratov napak na vzorcih — to pa v definiciji ohranja odvisnost od načina vzburjanja, na katerega velikokrat nimamo vpliva, kot je na primer področje identifikacije zaprtizančnih sistemov, kjer z regulacijskim signalom najprej zagotovimo kakovost vodenja, šele nato razmišljamo o signalu za dobro identifikacijo modela. V tem članku bomo prikazali primere, kjer VDF še vedno potrebuje vztrajno vzburjanje.

Zato v članku predlagamo algoritem za pozabljanje s spremenljivo smerjo na osnovi domene, ki se opira na pričakovano vrednost kvadratne napake nad pomembnim delom vhodne domene, s čimer odpravi odvisnost od vztrajnega vzburjanja. Glavni prispevki članka so:

- Izpostavitev pomanjkljivosti algoritma RLS-VDF.
- Razvoj novega algoritma, imenovanega RLS z domensko usmerjenim pozabljanjem (RLS-3DF), ki je izpeljan iz pričakovane vrednosti kvadrata napake in pozablja le tisti del kovariančne matrike, kjer lahko z upoštevanjem novega vzorca pričakujemo zmanjšanje pričakovane vrednosti kvadrata napake.
- Primerjava različnih rekurzivnih algoritmov na simuliranem statičnem sistemu.

2 Opis problema

Obravnavajmo časovno spremenljiv statični linearni proces z vhodnim vektorjem $\underline{x}^*$ dimenzije $(r_\phi - 1)$ in izhodom y^* , ki ga želimo identificirati v trenutku t , z izhodno enačbo:

$$y^*(t, \underline{x}^*) = \underline{\phi}^{*T}(t)\underline{\theta}(t) \quad (1)$$

kjer je $\underline{\phi}^*(t) = [\underline{\phi}_k^T(t), 1]^T$ regresijski vektor dimenzije r_ϕ , $\underline{\theta}(t)$ pa predstavlja realne parametre procesa v trenutku t .

Proces je vzorčen pri vsakem k -tem časovnem koraku t_k , kjer $\underline{x}_k$ predstavljata meritev vhodnega vektorja, y_k pa meritev izhoda. Meritev izhoda vnaša šum, medtem ko za namen članka predpostavimo nepošumljeno meritev vhodnega vektorja:

$$y_k = y^*(t_k) + e_y(t_k), \quad \underline{\phi}_k(t_k) = \underline{x}^*(t_k) \quad (2)$$

kjer je $e_y \sim \mathcal{N}\{0, \sigma_y^2\}$ normalno porazdeljen izhodni šum z varianco σ_y^2 .

Cilj je v vsakem časovnem koraku k oceniti realne parametre $\underline{\theta}$ z ocenjevalnikom $\hat{\underline{\theta}}$.

3 RLS-VDF

Eden izmed najpogostejše uporabljenih algoritmov za identifikacijo parametrov linearnih sistemov je rekurzivna metoda najmanjših kvadratov s pozabljanjem (RLS), ki je bila kasneje razširjena z variabilno-usmerjenim pozabljanjem (VDF). Oba pristopa sta predstavljena v tem razdelku; podrobnosti so podane v [9].

Pri metodi RLS ocenjujemo vektor parametrov $\underline{\theta}_k$ linearnega modela:

$$\hat{y}_k = \underline{\phi}_k^T \hat{\underline{\theta}}_k \quad (3)$$

kjer je $\hat{y}_k$ izhod modela v k -tem časovnem koraku.

Pri RLS z (skalarnim) faktorjem pozabljanja moramo rekurzivno izračunati dve spremenljivki pri koraku $k + 1$. Inverz kovariančne matrice $\underline{P}_{k+1}$:

$$\underline{P}_{k+1} = \frac{1}{\lambda} \underline{P}_k - \underline{P}_k \underline{\phi}_k (\underline{I}_p + \frac{1}{\lambda} \underline{\phi}_k^T \underline{P}_k \underline{\phi}_k)^{-1} \underline{\phi}_k^T \underline{P}_k \quad (4)$$

kjer je λ faktor pozabljanja, $\underline{I}_p$ pa enotska matrika velikosti r_ϕ . Ocenjeni parametri $\hat{\underline{\theta}}_{k+1}$:

$$\hat{\underline{\theta}}_{k+1} = \hat{\underline{\theta}}_k + \underline{P}_{k+1} \underline{\phi}_k (y_k - \underline{\phi}_k^T \hat{\underline{\theta}}_k) \quad (5)$$

3.1 Variabilno usmerjeno pozabljanje (VDF)

Ključna predpostavka algoritma RLS je vztrajno vzbujanje v vseh smereh. V nasprotnem primeru vsaj ena od singularnih vrednosti $\underline{P}_k$ divergira. Ena boljših rešitev je predstavljena v [9], kjer je opisan metoda VDF. Pri VDF je v enačbi (4) zamenjan skalar λ s podatkovno odvisno pozabljačo matriko $\underline{\Lambda}_k$:

$$\underline{P}_{k+1} = \bar{\underline{P}}_k - \bar{\underline{P}}_k \underline{\phi}_k (\underline{I}_p + \underline{\phi}_k^T \bar{\underline{P}}_k \underline{\phi}_k)^{-1} \underline{\phi}_k^T \bar{\underline{P}}_k \quad (6)$$

kjer

$$\bar{\underline{P}}_k = \underline{\Lambda}_k^{-1} \underline{P}_k \underline{\Lambda}_k^{-1} \quad (7)$$

Matrika $\underline{\Lambda}_k$ je pozitivno definitna (in torej simetrična), sestavljena pa je na naslednji način:

$$\bar{\underline{\Lambda}}_k(i, i) \triangleq \begin{cases} \sqrt{\lambda}, & \|\psi_{k,i}\| > \epsilon \\ 1, & \text{otherwise} \end{cases} \quad (8)$$

Najprej se na $\underline{P}_k^{-1}$ uporabi singularni razcep. Matrika $\underline{U}_k$ predstavlja ortonormalno matriko, katere stolpci so singularni vektorji matrike $\underline{P}_k^{-1}$. S pomočjo $\underline{U}_k$ se regresijski vektor $\underline{\phi}_k$ projicira na singularne vektorje:

$$\underline{\psi}_k = \underline{\phi}_k \underline{U}_k \quad (9)$$

Norma nastalega vektorja $\underline{\psi}_k$ kaže na informacijsko vsebino v $\underline{\phi}_k$ vzdolž singularnih vektorjev. S pomočjo $\underline{\psi}_k$ se sestavi diagonalna matrika $\bar{\underline{\Lambda}}$:

$$\bar{\underline{\Lambda}}_k(i, i) \triangleq \begin{cases} \sqrt{\lambda}, & \|\psi_{k,i}\| > \epsilon, \\ 1, & \text{drugače} \end{cases} \quad (10)$$

kjer je ϵ uporabniški prag za $\|\psi_{k,i}\|$, nad katerim se pozabljanje uporabi vzdolž ustreznega singularnega vektorja. Končna matrika $\underline{\Lambda}_k$ se sestavi na naslednji način:

$$\underline{\Lambda}_k = \underline{U}_k \bar{\underline{\Lambda}}_k \underline{U}_k^T \quad (11)$$

Rezultat lahko vidimo na sliki 1 — medtem ko v primeru regresijskega vektorja $\underline{\phi}_1$ pozabljam v vseh smereh, pa pri regresijskem vektorju $\underline{\phi}_2$ v smeri singularnega vektorja $\underline{u}_2$ ne vzbuja, saj je projekcija $\underline{\phi}_2$ na $\underline{u}_2$ manjša od praga ϵ .

3.2 Problemi RLS-VDF

Algoritem RLS-VDF ima dvoje pomanjkljivosti: vzbujanje v smereh, ki niso poravnane s singularnimi vektorji, ter uvedba dodatnega praga.

3.2.1 Vzbujanje v smereh, ki niso poravnane s singularnimi vektorji

Čeprav preverjanje vzbujanja v smereh singularnih vektorjev lahko omeji pozabljanje v nekaterih smereh, to ne odpravi vedno težave. Vzemimo primer iz [9], prikazan na sliki 1, in se osredotočimo na rdeč regresijski vektor $\underline{\phi}_1$. Po RLS-VDF pozabljam v vseh smereh, saj so vse projekcije večje od praga ϵ . Sedaj vzemimo hipotetičen primer, ko za nekaj časa vsi vzorci ležijo v smeri regresijskega vektorja $\underline{\phi}_1$. Kljub temu, da ni vzbujanja v smereh pravokotnega na $\underline{\phi}_1$, bo algoritem RLS-VDF še vedno pozabljal v vseh smereh. To lahko privede do singularne matrike, ali vsaj do pozabljanja pomembnega dela vzorcev. To je nezaželena posledica.

3.2.2 Uvedba praga

Čeprav nastavitev enega hiperparametra ni zahtevna, kadar obravnavamo RLS samostojno, postane v večjih sistemih z več modeli in regulacijskimi zankami vsak dodatni hiperparameter breme za umerjanje in razhroščevanje. Posledično lahko uvedba praga ϵ oteži načrtovanje kompleksnih sistemov.

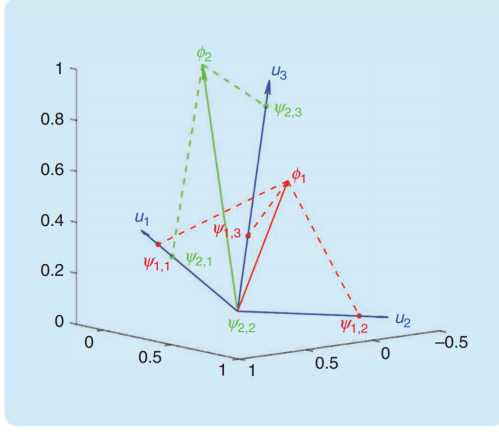


Figure 1: Modri vektorji predstavljajo singularne vektorje $\underline{U}_k$, rdeči (ϕ_1) in zeleni (ϕ_2) vektor predstavljata regresijski vektorja, s črtkano so pa označene projekcije regresijskih vektorjev na singularne vektorje. Povzeto iz [9].

4 Usmerjeno pozabljanje, vodeno z domeno, v k -tem koraku

Usmerjeno pozabljanje je izpeljano iz kriterijske funkcije kvalitete modela, za katero smo izbrali pričakovano vrednost kvadrata napake (ESE), namesto vsote kvadratov napak na razpoložljivih vzorcih (npr. pri RLS-VDF). Čeprav je ESE težje oceniti, nima težav z zahtevo po vztrajnem vzburjanju, saj je kriterijska funkcija neodvisna od vzorčenja. V tem razdelku bomo pokazali, da se lahko namesto vsote napak uporabi pričakovana vrednost kvadrata napake. Teorija pokaže, da je ob uporabi ESE zmanjšanje napake mogoče le v dveh smereh, tudi ob predpostavkah o pristranskosti parametrov.

Kriterijska funkcija je definirana kot ESE modela, kjer je regresijski vektor ϕ naključen, ki sledi neki verjetnostni porazdelitvi p_x nad vhodno domeno. Kriterijska funkcija ima torej obliko:

$$J(\hat{\theta}_k) = E \left\{ (f(\phi, k) - \phi^T \hat{\theta}_k)^2 \right\}, \quad \phi \sim p_x \quad (12)$$

kjer je $f(\phi, k)$ izhodna funkcija modeliranega sistema v času k , $\hat{\theta}_k$ pa ocenjeni vektor parametrov v istem trenutku.

Izpeljava je predstavljena v delu [14], v tem delu je predstavljen le rezultat.

Naj $\underline{c} = E \{ \phi \}$ predstavlja center dela domene, na kateri želimo točen model. V delu [14] je dokazano, da če matriko $\underline{P}_{k+1}$ pozabimo v kateri koli smeri, ki ni poravnana s $\underline{P}_{k+1}\underline{c}$ ali $\underline{P}_{k+1}(\phi_k - \underline{c})$, bo vsak vnos pozabljanja λ povzročil zgolj povečanje kriterijske funkcije $J(\hat{\theta}_{k+1})$. Torej pri uporabi faktorja pozabljanja se lahko kriterijska funkcija zmanjša le v smeri $\underline{P}_{k+1}\underline{c}$ ali $\underline{P}_{k+1}(\phi_k - \underline{c})$. S tem smo določili edine smiselne smeri pozabljanja.

5 RLS s pozabljanjem z variabilno smerjo na osnovi domene (RLS-3DF)

Namesto pozabljanja celotne korelacijske matrike, ki združuje središče in kovariančno matriko, v tem

članku predlagamo razcep korelacijske matrike $\underline{P}_k^{-1}$ na kovariančni člen in člen središča:

$$\underline{P}_k^{-1} = \underline{G}_k + s_k \underline{c}_k \underline{c}_k^T \quad (13)$$

kjer s_k ustreza številu vzorcev. Ta razcep omogoča ločeno pozabljanje člena središča $s_k \underline{c}_k \underline{c}_k^T$ (s tem pozabljanjem v smeri $\underline{P}_{k+1}\underline{c}$) in kovariančnega člena $\underline{G}_k$ (ki omogoča pozabljanje v smeri $\underline{P}_{k+1}(\phi_k - \underline{c})$).

Takšna definicija ima za posledico dodatno rekurzivno izračunavanje $\underline{c}_k$ in s_k — podrobneje je razloženo v naslednjem razdelku.

Algoritem RLS-3DF je sestavljen iz dveh korakov:

1. Pozabljanje $\underline{P}_k^{-1}$ preko $\underline{G}_k$ in s_k .
2. Posodobitev $\underline{P}_{k+1}^{-1}$, $\hat{\theta}_{k+1}$, $\underline{c}_{k+1}$, s_{k+1} v koraku $k+1$.

5.1 Pozabljanje $\underline{G}_k$ in s_k v k -tem koraku

Najprej izračunamo kovariančno matriko $\underline{G}_k$, izraženo iz (13):

$$\underline{G}_k = \underline{P}_k^{-1} - s_k \underline{c}_k \underline{c}_k^T \quad (14)$$

Predlagamo, da pozabimo $\underline{G}_k$ v smeri vektorja od središča $\underline{c}_k$ do regresijskega vektorja ϕ_k , uteženega s $\underline{G}_k$. Smer določa enotski vektor $\underline{n}_{c,k}$:

$$\underline{n}_{c,k} = \frac{\underline{G}_k \underline{d}}{\|\underline{G}_k \underline{d}\|} \quad (15)$$

kjer $\underline{d} = \phi_k - \underline{c}_k$. Nato uvedemo matriko pozabljanja kovariančne matrike $\underline{\Lambda}_{c,k}$, s katero omogočimo pozabljanje le v smeri $\underline{n}_{c,k}$:

$$\underline{\Lambda}_{c,k} = \underline{I} - (1 - \sqrt{\lambda}) \underline{n}_{c,k} \underline{n}_{c,k}^T \quad (16)$$

kjer λ predstavlja faktor pozabljanja. Matriko $\underline{G}_k$ pozabimo v smeri $\underline{n}_{c,k}$ z enačbo:

$$\bar{\underline{G}}_k = \underline{\Lambda}_{c,k} \underline{G}_k \underline{\Lambda}_{c,k} \quad (17)$$

s čimer dobimo posodobljeno kovariančno matriko $\bar{\underline{G}}_k$.

Sedaj pozabimo še v smeri $\underline{c}$, kar dosežemo s pozabljanjem števila vzorcev:

$$\bar{s}_k = \lambda s_k \quad (18)$$

Opozoriti velja, da v tem članku uporabljamo enak faktor pozabljanja tako za središče kot za $\underline{G}_k \underline{d}$ za čistejši zapis pozabljanja, vendar ne obstaja nobena omejitev zakaj razvijalec ne bi nastavljal različnih vrednosti λ za vsako komponento posebej.

Korelacijska matrika po izvedenem pozabljanju, označena kot $\bar{\underline{P}}$, se rekonstruira iz izrazov (17) in (18):

$$\bar{\underline{P}}_k^{-1} = \bar{\underline{G}}_k + \bar{s}_k \underline{c}_k \underline{c}_k^T \quad (19)$$

5.2 Posodabljanje v koraku $k+1$

V koraku $k+1$ posodobimo vse potrebne vrednosti $\underline{P}_{k+1}^{-1}$, $\hat{\theta}_{k+1}$, $\underline{c}_{k+1}$, s_{k+1} . Najprej posodobimo korelacijsko matriko $\underline{P}_{k+1}^{-1}$:

$$\underline{P}_{k+1}^{-1} = \bar{\underline{P}}_k^{-1} + \phi_k \phi_k^T \quad (20)$$

Na podlagi P_{k+1}^{-1} lahko ocenimo vektor parametrov $\hat{\theta}_{k+1}$ z enako enačbo kot pri metodi RLS (glej (5)):

$$\hat{\theta}_{k+1} = \hat{\theta}_k + P_{k+1} \phi_k (y_k - \phi_k^T \hat{\theta}_k) \quad (21)$$

Opozoriti velja, da je P_{k+1} treba izračunati neposredno z inverzijo matrike P_{k+1}^{-1} . Če so se raziskovalci v zgodovini takih inverzov izogibali, je naše mnenje, da imamo dandanes že tako hitro strojno opremo, da izračun inverza ne predstavlja prevelikega problema. Kljub temu lahko numerična stabilnost direktnega inverza še vedno predstavlja izziv, zato spodbujamo druge raziskovalce, da izpeljejo rekurzivno enačbo za P_{k+1} .

Za naslednji korak posodabljanja je treba posodobiti vektor središča $\underline{c}_{k+1}$ in utež središča s_{k+1} . Uporabimo pravila za posodabljanje iz rekurzivnega Gustafson-Kessel algoritma za Takagi-Sugeno modele [15]:

$$\underline{c}_{k+1} = \underline{c}_k + \bar{s}_k^{-1} (\underline{x}_k - \underline{c}_k) \quad (22)$$

$$s_{k+1} = \bar{s}_k + 1 \quad (23)$$

6 Eksperimenti in rezultati

6.1 Eksperiment

Rekurzivni algoritmi so bili testirani v dveh scenarijih:

- (E1) kadar je sistem vztrajno vzbujen — v tem primeru naj bi se algoritem obnašal podobno kot pri standardnem RLS.
- (E2) kadar sistem ni vztrajno vzbujen.

6.1.1 Časovno spremenljiv statični sistem (TVSS)

Modelirani sistem je časovno spremenljiv linearni model z dodanim šumom ($e_y \sim \mathcal{N}\{0, 2\}$):

$$y(t) = \phi^T(t) \theta(t) + e_y \quad (24)$$

Model mora biti natančen znotraj vzorčevalnega območja:

$$\underline{x} \sim \mathcal{N}\{\underline{c}, \underline{A}\} \quad (25)$$

kjer sta:

$$\underline{c} = \begin{bmatrix} 5 \\ 3 \\ 2 \\ 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad \underline{A} = \begin{bmatrix} 0.4171 & & & & & & \\ & 1.3675 & & & & & \\ & & 1.5614 & & & & \\ & & & 1.8113 & & & \\ & & & & 0.0676 & & \\ & & & & & 1 & \\ & & & & & & \end{bmatrix} \quad (26)$$

Vektor parametrov $\theta(t)$ je bil izbran po izrazu (27):

$$\theta(t) = \theta_{pre}(t) - \begin{bmatrix} 0_{r_\phi-1} \\ \theta_{pre}^T(t) [c_1, \dots, c_{r_\phi-1}, 0]^T \end{bmatrix} \quad (27)$$

kjer velja:

$$\theta_{pre}(t) = \theta_0 + 2 \cos \left(\begin{bmatrix} \frac{t-200}{100} \\ \frac{t-200}{90} \\ \frac{t-200}{80} \\ \frac{t-200}{70} \\ \frac{t-200}{60} \\ \frac{t-200}{20} \end{bmatrix} \right), \quad \theta_0 = \begin{bmatrix} 1 \\ 10 \\ 1 \\ -2 \\ 1 \\ 6 \end{bmatrix} \quad (28)$$

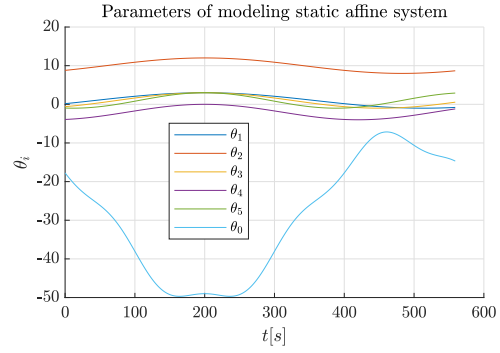


Figure 2: Vrednosti parametrov TVSS

Vektor parametrov je bil zasnovan tako, da se sčasoma počasi odmika v različnih smereh, kar omogoča preizkus sposobnosti algoritmov za modeliranje časovno spremenljivih procesov.

6.1.2 Podatkovna množica

Časovno spremenljivi statični sistem (TVSS) je bil zaporedno vzorčen z uporabo petih različnih normalnih porazdelitev, vse s središčem v $\underline{c}_i = \underline{c}$. Tako smo dobili pet skupin vzorcev $\mathcal{S}_i$, $i \in 1, \dots, 5$.

Prva skupina ($N_1 = 200$ vzorcev) je bila vztrajno vzbujena ($\underline{A}_1 = \underline{A}$). Napaka v tej skupini ustreza sposobnosti algoritma, da modelira sisteme z vztrajnim vzbujanjem.

Po $\mathcal{S}_1$ so sledile tri skupine z nezadostnim vzbujanjem, vsaka s $N_2 = N_3 = N_4 = 20$ vzorci. Te skupine so imele rang-ena kovariančne matrike, kjer je bil lastni vektor, ki pripada neničelni lastni vrednosti, pravokoten na tiste iz ostalih rang-ena kovariančnih matrik. Takšna nastavitve omogoča oceno vpliva singularnega vzorčenja na ocenjevanje parametrov.

Zadnja skupina ($N_5 = 200$ vzorcev) je bila ponovno vztrajno vzbujena ($\underline{A}_5 = \underline{A}$), kar omogoča oceno, koliko informacij se je izgubilo zaradi nezadostnega vzbujanja v predhodnih skupinah.

Povzetek matrik $\underline{A}_i$ in velikosti vzorčenja N_i za vsako skupino i je prikazan v naslednji tabeli:

Group i	Covariance Matrix $\underline{A}_i$	N_i
$\mathcal{S}_1$	$\underline{A}_1 = \underline{A}$	200
$\mathcal{S}_2$	$\underline{A}_2 = \frac{a_2 a_2^T}{\sqrt{a_2^T \underline{A} a_2}}$	20
$\mathcal{S}_3$	$\underline{A}_3 = \frac{a_3 a_3^T}{\sqrt{a_3^T \underline{A} a_3}}$	20
$\mathcal{S}_4$	$\underline{A}_4 = \frac{a_4 a_4^T}{\sqrt{a_4^T \underline{A} a_4}}$	20
$\mathcal{S}_5$	$\underline{A}_5 = \underline{A}$	300

kjer so

$$\begin{aligned} a_2 &= \frac{1}{\sqrt{5}} [1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 0]^T \\ a_3 &= \frac{1}{\sqrt{5}} [1 \quad -1 \quad 0 \quad -1 \quad 1 \quad 0]^T \\ a_4 &= \frac{1}{\sqrt{5}} [1 \quad 0 \quad 0 \quad 0 \quad -1 \quad 0]^T \end{aligned}$$

6.2 Testirani algoritmi

Testirani algoritmi so bili klasični RLS, RL VDF in predlagani RLS-3DF. Vsi algoritmi bili ocenjeni z 41 različnimi faktorji pozabljanja $\lambda \in \{0.2, 0.22, 0.24, \dots, 1\}$, pri čemer je bila za očejevanje izbrana optimalna vrednost glede na doseženo učinkovitost.

Merilo učinkovitosti je koren povprečnega kvadratne napake (RMSE):

$$RMSE_i = \sqrt{\frac{\sum_{k \in S_i} (\hat{y}(k) - y(k))^2}{N_i}}, \quad (2)$$

6.3 Rezultati

V tem razdelku so predstavljeni rezultati eksperimento

6.3.1 Evalvacija rezultatov RMSE

Rezultati RMSE za posamezne skupine vzorcev so povzeti v tabeli 1. Čeprav smo pričakovali, da bo RLS-3DF izboljšal RMSE le v primerih nezadostnega vzbujanja (skupine 2–5), je presenetljiva ugotovitev, da se je RMSE zmanjšal tudi v pogojih vztrajnega vzbujanja. To poudarja prednost prehoda s kriterijskih funkcij, temelječih na vzorcih, na kriterijsko funkcijo, podano s pričakovano vrednostjo kvadrata napake (ESE).

Table 1: Primerjava RMSE različnih skupin

	RLS	RLS-VDF	RLS-3DF	Izboljšava[%]
1	7.87	7.63	5.12	32.9
2	25.97	25.27	17.87	29.3
3	26.16	25.48	17.97	29.5
4	26.67	25.77	18.33	28.9
5	9.70	9.42	7.33	22.2

6.3.2 Najboljši faktorji pozabljanja

Naslednji zanimiv rezultat se nanaša na izbiro najboljših faktorjev pozabljanja za posamezne algoritme. Na sliki 3 je prikazan graf $\log_2(\text{RMSE})$ glede na λ za skupini S_1 in S_4 . Prikaz z $\log_2$ je bil izbran zaradi tega, ker omogoča hkraten prikaz minimumov ter eksponentnega naraščanja RMSE - drugače bi morali izbirati ali naj slika prikazuje relacije med najboljšimi vrednostmi ali potek naraščanja. Poleg očitno nižje vrednosti RMSE pri algoritmu RLS-3DF v primerjavi z drugima dvema algoritmoma bi izpostavili še dve presenetljivi ugotovitvi:

Ključna ugotovitev je, da je bil optimalni faktor pozabljanja za RLS-3DF bistveno nižji, kot smo pričakovali — okoli $\lambda \sim 0,65$, tudi v regresijskem prostoru s šestimi dimenzijami. To omogoča hitrejšo konvergenco parametrov, kar je ključno pri modeliranju hitro spreminjajočih se sistemov. Razlog za to je verjetno v tem, da v vsakem koraku k pozabljamo le v dveh smereh.

Poleg tega ima funkcija RMSE pri algoritmu RLS-3DF bistveno manjša nihanja glede na λ kot pri drugih algoritmihih. Zaradi tega je RLS-3DF primernejši za postopke optimizacije, ki optimizirajo λ (boljša ocena gradienta), in tudi lažji za uporabnike pri določanju ustreznega faktorja pozabljanja.

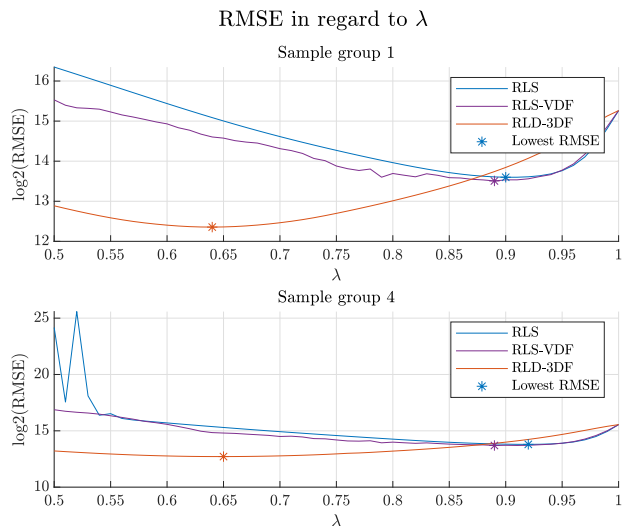


Figure 3: RMSE v odvisnosti od vrednosti λ

7 Zaključek

V članku smo predstavili nov rekurzivni algoritem najmanjših kvadratov (RLS) z usmerjenim pozabljanjem, ki temelji na vhodni domeni (RLS-3DF). Predlagani pristop odpravi ključne pomanjkljivosti klasičnega algoritma z variabilnim usmerjenim pozabljanjem (RLS-VDF). Predlagani pristop z uporabo kriterijske funkcije, podane s pričakovano vrednostjo kvadrata napake (ESE), odpravlja pogoj vztrajnega vzbujanja. Tako je kriterijska funkcija definirana izključno v vhodni domeni, namesto da bi se zanašal na vzorčene podatke in posledično pravilno vzorčenje. Ločena posodobitev kovariančne matrike in središča modela RLS-3DF omogoča učinkovitejšo pozabljanje, zlasti v slabo vzbujenih smereh, ob tem pa še vedno ohranja zanesljivo ocenjevanje parametrov.

Eksperimentalni rezultati so pokazali, da RLS-3DF presega tako standardni RLS kot RLS-VDF v vseh testiranih primerih, vključno z vztrajno in nezadostno vzbujenimi pogoji. Predlagani algoritem je znižal RMSE za vsaj 20 % na vseh testnih primerih, kar potrjuje njegovo boljšo oceno parametrov. Poleg tega je RLS-3DF izkazal nižjo občutljivost na faktor pozabljanja, kar poenostavi njegovo nastavitve v praktičnih aplikacijah.

Prihodnje raziskave bodo usmerjene v razširitev algoritma RLS-3DF z uporabo variabilnega faktorja pozabljanja. Predvidena je njegova uporaba pri identifikaciji nelinearnih sistemov in v adaptivnih regulacijskih aplikacijah. Potrebna pa bo tudi nadaljnja analiza teoretičnih lastnosti konvergence ter robustnosti metode v visokodimenzionalnih regresijskih prostorih.

Zahvala

Delo je bilo izvedeno v okviru programa DIGITOP, financiran s strani Ministrstva za visoko šolstvo, znanost in inovacije Republike Slovenije, Javne agencije za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije ter Evropske Unije.

References

- [1] G. Andonovski, E. Lughofer, and I. Škrjanc, “Evolving Fuzzy Model Identification of Nonlinear Wiener-Hammerstein Processes,” *IEEE Access*, vol. 9, pp. 158 470–158 480, 2021.
- [2] G. Karer, I. Škrjanc, and B. Zupančič, “Self-adaptive predictive functional control of the temperature in an exothermic batch reactor,” *Chemical Engineering and Processing: Process Intensification*, vol. 47, no. 12, pp. 2379–2385, 2008.
- [3] G. Andonovski, B. S. J. Costa, S. Blažič, and I. Škrjanc, “Robust evolving controller for simulated surge tank and for real two-tank plant,” *at - Automatisierungstechnik*, vol. 66, no. 9, pp. 725–734, Sep. 2018.
- [4] M. Ozbot, E. Lughofer, and I. Škrjanc, “Evolving Neuro-Fuzzy Systems based Design of Experiments in Process Identification,” *IEEE Transactions on Fuzzy Systems*, pp. 1–11, 2022.
- [5] G. Andonovski, G. Mušič, and I. Škrjanc, “Fault detection through evolving fuzzy cloud-based model,” *IFAC-PapersOnLine*, vol. 51, no. 2, pp. 795–800, 2018.
- [6] M. Pratama, E. Dimla, T. Tjahjowidodo, W. Pedrycz, and E. Lughofer, “Online Tool Condition Monitoring Based on Parsimonious Ensemble+,” *IEEE Transactions on Cybernetics*, vol. 50, no. 2, pp. 664–677, Feb. 2020.
- [7] C. Paleologu, J. Benesty, and S. Ciochiña, “A robust variable forgetting factor recursive least-squares algorithm for system identification,” *IEEE Signal Processing Letters*, vol. 15, no. 3, pp. 597–600, 2008.
- [8] G. Černe and I. Škrjanc, “Cascade affine constant recursive algorithm for model-based control,” *IFAC-PapersOnLine*, vol. 54, no. 14, pp. 299–302, 2021.
- [9] A. Goel, A. L. Bruce, and D. S. Bernstein, “Recursive Least Squares With Variable-Direction Forgetting: Compensating for the Loss of Persistence [Lecture Notes],” *IEEE Control Systems*, vol. 40, no. 4, pp. 80–102, Aug. 2020.
- [10] M. Versaci and F. La Foresta, “EEG Data Analysis Techniques for Precision Removal and Enhanced Alzheimer’s Diagnosis: Focusing on Fuzzy and Intuitionistic Fuzzy Logic Techniques,” *Signals*, vol. 5, no. 2, pp. 343–381, Jun. 2024.
- [11] K. Zhang, S. Luo, H.-N. Wu, and R. Su, “Data-Driven Tracking Control for Nonaffine Yaw Channel of Helicopter via Off-Policy Reinforcement Learning,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 61, no. 3, pp. 7725–7737, Jun. 2025.
- [12] W. T. Chor, C. P. Tan, A. S. M. Bakibillah, Z. Pu, and J. Y. Loo, “Robust Vehicle Mass Estimation Using Recursive Least M-Squares Algorithm for Intelligent Vehicles,” *IEEE Transactions on Intelligent Vehicles*, vol. 9, no. 1, pp. 165–177, Jan. 2024.
- [13] M. Al Ahdab, T. Knudsen, J. Stoustrup, and J. Leth, “An Online Stochastic Optimization Approach for Insulin Intensification in Type 2 Diabetes with Attention to Pseudo-Hypoglycemia,” in *2022 IEEE 61st Conference on Decision and Control (CDC)*. Cancun, Mexico: IEEE, Dec. 2022, pp. 2572–2579.
- [14] G. Černe and I. Škrjanc, “Recursive least squares with domain-driven variable-direction forgetting,” in *2025 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. Reims, France: IEEE, Jul. 2025, pp. 1–6.
- [15] D. Dovžan and I. Škrjanc, “Recursive clustering based on a Gustafson-Kessel algorithm,” *Evolving Systems*, vol. 2, no. 1, pp. 15–24, 2011.