Lekcija 2: *On-line estimacija parametara*

Prof.dr.sc. Jasmin Velagić
Elektrotehnički fakultet Sarajevo

Kolegij: Adaptivno i robusno upravljanje

2012/2013

Osnovni pojmovi o estimaciji parametara

- **Estimacija je proces procjene vrijednosti veličina od interesa na temelju dostupnih mjerenja, koja su uglavnom posredna, netačna i/ili neizvjesna (nesigurna).**
- Estimaciju su prvi koristili Laplace, Legandre i Gauss za određivanje orbitalnih parametara planete.
- Estimacija se može koristiti u mnogim područjima:
  - ➢ Statističko zaključivanje (izbor najboljeg rješenja iz diskretnog skupa mogućih rješenja).
  - ➢ Određivanje položaja i brzine pokretnih objekata – slijeđenje (tracking).
    - ✓ Primjer: estimacija stanja pokretnog objekta na temelju mjerenja dobivenih iz udaljenih senzora, postavljenih na fiksnim lokacijama ili pokretnim platformama.

Osnovni pojmovi o estimaciji parametara

- Upravljanje sistemima uz prisustvo neodređenosti (šum i/ili nepoznati parametri) – identifikacija parametara, estimacija stanja i stohastičko upravljanje.
- Određivanje parametara matematičkog modela radi predviđanja stanja stvarnog sistema – identifikacija sistema.
  - ✓ Primjeri: estimacija parametara leta aviona, estimacija parametara rotorskog kruga asinhronog motora, estimacija parametara hoda nožnog robota, estimacija parametara kontakta kotača i podloge,...
- Određivanje svojstava prenesene poruke na temelju zašumljenog primljenog signala – teorija komunikacija.

Osnovni pojmovi o estimaciji parametara

- Određivanje nekih parametara ili značajki signala ili slike – obrada signala/slike, ...

- Gauss je u svojim istraživanjima došao do sljedećih zaključaka o estimaciji:
  - **Matematički model sistema je dostupan, ali su neki parametri nepoznati i treba ih estimirati.**
  - **Neophodni su redundantni podaci za smanjenje učinaka pogrešaka mjerenja.**
  - **Rezidui trebaju biti što je moguće manji.**
  - **Netačnosti mjerenja zahtijevaju primjenu pristupa modeliranja zasnovanog na teoriji vjerojatnosti.**
  - **Kombinacija početnog znanja i slijeda naknadnih mjerenja vodi ka primjeni koncepta rekurzivnih algoritama.**

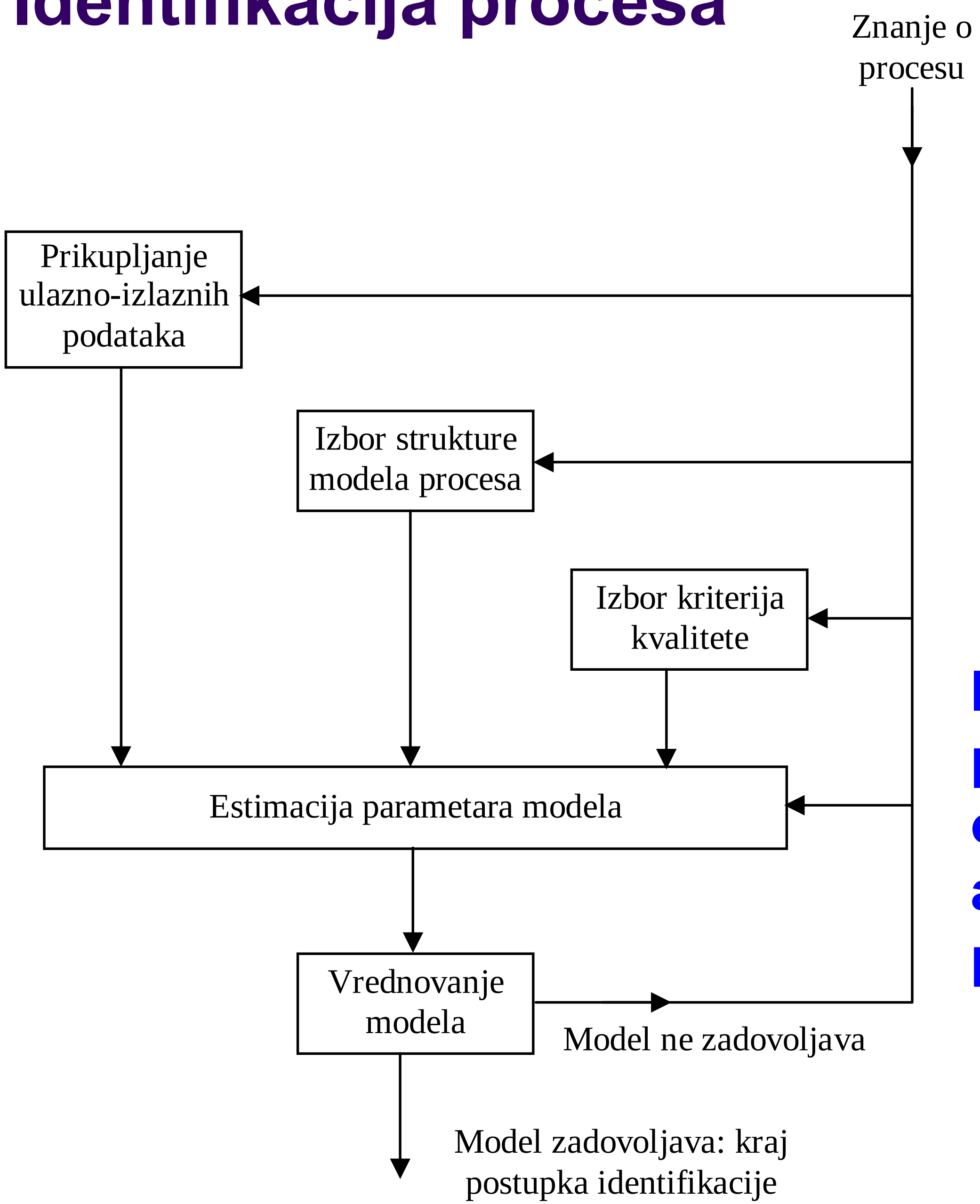
Osnovni pojmovi o estimaciji parametara

- **Optimalni estimator je algoritam koji estimira promatranu veličinu obrađujući dostupna mjerenja i optimirajući određeni kriterij.**
- **Prednosti** korištenja optimalnog estimatora: najbolje iskorištenje podataka i znanja o sistemu i poremećajima na sistem.
- **Nedostaci**: moguća osjetljivost na pogreške modeliranja, moguća velika računarska složenost.
- Veličine (varijable) koje se estimiraju:
  - **Parametar** - vremenski nepromjenjiva (ili sporo promjenjiva) veličina (skalar, vektor ili matrica).
  - **Stanje** – vremenski promjenjiva veličina dinamičkog stanja (obično vektor).
- Estimatori se dijele na: estimatore parametara (identifikacija parametara sistema) i estimatore stanja.

Identifikacija procesa

- **Postupak identifikacije procesa**
- **Eksperimentalno matematičko modelliranje sistema na temelju skupa izmjerenih vrijednosti ulaznih i izlaznih signala sistema naziva se *identifikacijom sistema*.**
- Postupak identifikacije procesa odvija se u nekoliko osnovnih koraka:
  - **Izbor strukture modela procesa;**
  - **Prikupljanje ulazno-izlaznih podataka, tj. mjerenje vrijednosti ulaznih i izlaznih signala procesa;**
  - **Izbor kriterija kvalitete modela procesa;**
  - **Estimacija parametara modela procesa;**
  - **Izbor optimalne dimenzije modela i njegovo vrednovanje.**

Identifikacija procesa

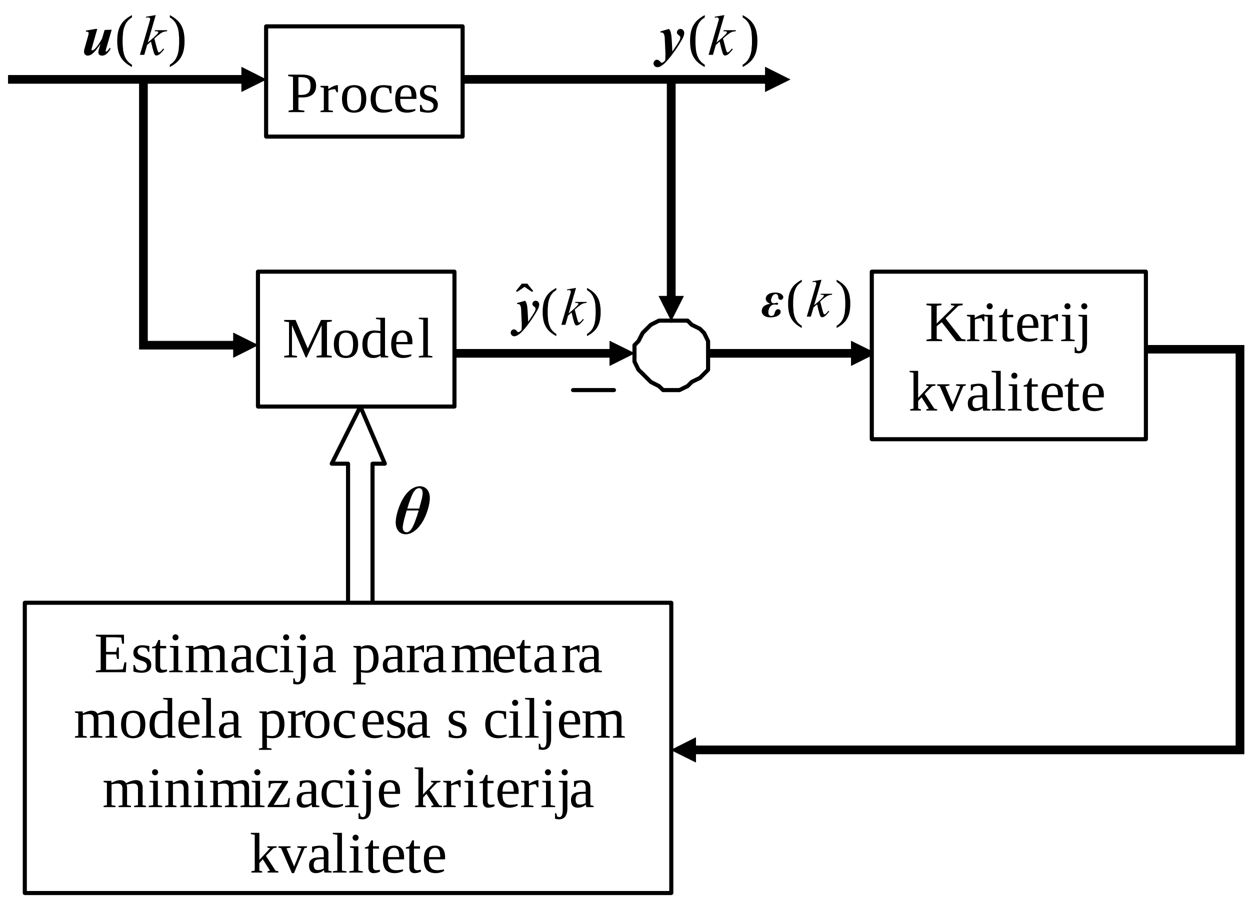


Dijagram toka postupka identifikacije procesa

Identifikacija sistema – potrebno je prvo odabrati model sistema, a nakon toga estimirati parametre modela.

Identifikacija procesa

- **Tok postupka identifikacije**
- Postupak identifikacije procesa odvija se u nekoliko koraka.



Identifikacija procesa

- Dovođenje pobudnih signala $\boldsymbol{u}(k)$ na ulaz procesa i modela izabrane strukture i početnih vrijednosti parametara.
- Računanje izlaznih signala hipotetičkog modela sa inicijalno pretpostavljenim vrijednostima parametara, označenih sa $\hat{\boldsymbol{y}}(k)$.
- Upoređivanje $\hat{\boldsymbol{y}}(k)$ s mjerenim vrijednostima izlaznih signala procesa $\boldsymbol{y}(k)$, iz čega se računa signal pogreške $\boldsymbol{\varepsilon}(k)$.
- Parametar k označava vremenski trenutak $t = kT$, gdje je T konstantan period uzorkovanja.
- Iz signala pogreške računa se iznos kriterija kvalitete (performance criterion ili performance index), koji predstavlja mjeru veličine pogreške.

Identifikacija procesa

- **Iznos kriterija kvaliteta iskazuje ovisnost pogreške o parametrima modela procesa Θ.**
- Zatim se nekim od **iterativnih numeričkih postupaka minimizacije estimiraju vrijednosti parametara modela koje minimiziraju kriterij kvalitete**.
- Na kraju postupka **dobiveni model procesa se podvrgava testu vrednovanja**.
- Ponekad se umjesto jednog modela identificiraju parametri grupe modela, gdje se najprije izabire najbolji model, a zatim se on podvrgava testu vrednovanja.
- Najčešće model procesa dobiven u prvoj iteraciji ne prolazi test vrednovanja, pa se neki koraci postupka identifikacije moraju ponoviti više puta, zbog čega se postupak identifikacije gotovo uvijek provodi iterativno.

Identifikacija procesa

- Opis nelinearnog dinamičkog procesa diskretnim jednadžbama u prostoru stanja:

$$\boxed{\boldsymbol{x}(k+1)=\boldsymbol{g}(k,\boldsymbol{x}(k),\boldsymbol{u}(k),\boldsymbol{w}(k))}$$

$$\boxed{\boldsymbol{y}(k)=\boldsymbol{h}(k,\boldsymbol{x}(k),\boldsymbol{u}(k),\boldsymbol{v}(k))} \qquad (1)$$

$\boldsymbol{u}(k)$ - vektor ulaznih signala procesa;
$\boldsymbol{x}(k)$ - vektor varijabli stanja procesa;
$\boldsymbol{y}(k)$ - vektor izlaznih signala procesa;
$\boldsymbol{g}$ - vektorska funkcija koja opisuje dinamiku procesa;
$\boldsymbol{w}(k)$ - vektor slučajnih varijabli, tzv. procesni šum;
$\boldsymbol{v}(k)$ - vektor slučajnih varijabli, tzv. mjerni šum;
$\boldsymbol{h}$ - vektorska funkcija koja opisuje ovisnost izlaznih signala procesa o varijablama stanja, ulazima i smetnjama.

Identifikacija procesa

- Od primarne važnosti je poznavanje funkcije koja opisuje ulazno-izlazno ponašanje.
- Jedan od načina prikaza je:

$$\boxed{\boldsymbol{y}(k) = \boldsymbol{f}(k, \boldsymbol{u}^{k-1}, \boldsymbol{y}^{k-1}) + \boldsymbol{v}(k)} \quad (2)$$

gdje su:

$\boldsymbol{f}$ - vektorska funkcija koja opisuje ovisnost izlaznih signala procesa o ulaznim signalima;

$$\boldsymbol{u}^{k-1} = [\boldsymbol{u}(1), \ldots, \boldsymbol{u}(k-1)]^{\boldsymbol{T}} = [u_1(1), \ldots, u_n(1), \ldots, u_1(k-1), \ldots, u_n(k-1)]^{\boldsymbol{T}}$$

matrica dostupnih mjernih vrijednosti ulaznih signala procesa u $(k$-$1)$ -tom trenutku, dimenzije $(k$-$1)\times n$;

$$\boldsymbol{y}^{k-1} = [\boldsymbol{y}(1), \ldots, \boldsymbol{y}(k-1)]^{\boldsymbol{T}} = [y_1(1), \ldots, y_n(1), \ldots, y_1(k-1), \ldots, y_n(k-1)]^{\boldsymbol{T}}$$

Identifikacija procesa

$\boldsymbol{y}^{k-1}$ matrica dostupnih mjernih vrijednosti izlaznih signala procesa u $(k$-$1)$-tom trenutku, dimenzije $(k$-$1)\times n$;

$$\boldsymbol{v}^{k-1}=\left[\boldsymbol{v}(1),\ldots,\boldsymbol{v}(k-1)\right]^{\boldsymbol{T}}=\left[v_1(1),\ldots,v_n(1),\ldots,v_1(k-1),\ldots,v_n(k-1)\right]^{\boldsymbol{T}}$$

matrica dostupnih mjernih vrijednosti mjernog šuma u $(k$-$1)$-tom trenutku, dimenzije $(k$-$1)\times n$;

- Prvi član jednadžbe (2) je funkcija prošlih mjernih vrijednosti ulaznih i izlaznih signala procesa, drugi član je neovisan o njima pa se ne može niti identificirati.
- **Problem identifikacije se svodi na pronalaženje aproksimacijske funkcije funkciji f.**

Identifikacija procesa

- Kao aproksimacijska funkcija uobičajeno se primjenjuje funkcija definirana vektorom parametara $\boldsymbol{\Theta}$ sa konačnim dimenzijama:

$$\boxed{\boldsymbol{f}_N(k, \boldsymbol{u}^{k-1}, \boldsymbol{y}^{k-1}, \boldsymbol{\Theta})} \tag{3}$$

- Parametri funkcije $\boldsymbol{f}_N$ određuju strukturu modela procesa:

$$\boxed{\hat{\boldsymbol{y}}(k) = \boldsymbol{f}_N(k, \boldsymbol{u}^{k-1}, \boldsymbol{y}^{k-1}, \boldsymbol{\Theta})} \tag{4}$$

- Vektor izlaznih signala modela procesa $\hat{\boldsymbol{y}}(k)$ izračunava se u $(k$-$1)$-om koraku, na osnovu dostupnih mjernih vrijednosti ulaznih i izlaznih signala procesa $[\boldsymbol{u}^{k-1}, \boldsymbol{y}^{k-1}]$ u tom trenutku.

Identifikacija procesa

- Vektor $\hat{\boldsymbol{y}}(k)$ predstavlja procijenjenu vrijednost vektora izlaznih signala procesa u k-tom koraku $\boldsymbol{y}(k)$ izračunatu jedan korak unaprijed, u $(k\text{-}1)$-om koraku.
- Model (4) naziva se ***predikcijski model procesa*** ili jednostavno ***prediktor***, a vektor signala pogreški između izlaznih signala procesa i modela dan sa:

$$\boxed{\boldsymbol{\varepsilon}(k) = \boldsymbol{y}(k) - \hat{\boldsymbol{y}}(k)}$$

  naziva se ***vektorom predikcijskih pogrešaka***.
- Nakon što je odabrana struktura modela (4) potrebno je estimirati vrijednosti parametara modela $\boldsymbol{\Theta} = \boldsymbol{\Theta}^*$ uz koje je predikcijska pogreška najmanja, odnosno uz koje izlaz modela predstavlja očekivanje izlaza procesa:

$$\boxed{\hat{\boldsymbol{y}}(k) = \boldsymbol{f}_N(k, \boldsymbol{u}^{k-1}, \boldsymbol{y}^{k-1}, \boldsymbol{\Theta}^*) = \boldsymbol{E}\{\boldsymbol{y}(k)\}}$$

Identifikacija procesa

- **Kada je zadan kriterij kvalitete, optimalne vrijednosti parametara modela se estimiraju odgovarajućim numeričkim postupcima.**
- Estimacija koristi skup mjernih podataka ulaznih i izlaznih signala procesa $r = 1,\ldots, N$ prikupljenih eksperimentom.
- **Vektor predikcijske pogreške** za r-ti vektor mjernih podataka označen je sa $\varepsilon(r,\boldsymbol{\Theta})$.
- Model procesa može se smatrati dobrim tek kada ukupna predikcijska pogreška $\varepsilon^*(\boldsymbol{\Theta})$ na čitavom skupu mjernih podataka poprimi najmanji iznos (off-line identifikacija).

Identifikacija procesa

- **Ukupna predikcijska pogreška** može se prikazati kao matrica dimenzije $n \times N$ koja se dobije slaganjem vektora $\boldsymbol{\varepsilon}(r, \boldsymbol{\Theta})$:

$$
\begin{aligned}
\boldsymbol{\varepsilon}^{*}(\boldsymbol{\Theta}) &= [\boldsymbol{\varepsilon}^{T}(1, \boldsymbol{\Theta}), \ldots, \boldsymbol{\varepsilon}^{T}(N, \boldsymbol{\Theta})] \\
&= \left[\varepsilon_{1}(1, \boldsymbol{\Theta}), \ldots, \varepsilon_{n}(1, \boldsymbol{\Theta}), \ldots, \varepsilon_{1}(N, \boldsymbol{\Theta}), \ldots, \varepsilon_{n}(N, \boldsymbol{\Theta})\right]
\end{aligned}
\tag{5}
$$

- **Iznos ukupne predikcijske pogreške** mjeri se kriterijem kvalitete, koji može biti bilo koja norma u prostoru vektora (5).
- **Najčešće je kriterij kvaliteta definiran normom oblika**:

$$
J(\boldsymbol{\Theta}) = \sum_{r=l}^{N} J_{r}(\boldsymbol{\Theta}) = \sum_{r=l}^{N} l(\boldsymbol{\varepsilon}(r, \boldsymbol{\Theta}))
$$

Identifikacija procesa

- U prethodnom izrazu su:

$J_r(\boldsymbol{\Theta})$ - iznos kriterijske funkcije na *r*-tom mjernom uzorku, tzv. lokalna funkcija gubitaka;

$l(\cdot)$ - skalarna, obično pozitivna, funkcija.

- Standardno se za funkciju $l(\cdot)$ uzima kvadratna norma:

$$\boxed{l(\boldsymbol{\varepsilon}(r,\boldsymbol{\Theta})) = \frac{1}{2}\boldsymbol{\varepsilon}^2(r,\boldsymbol{\Theta})}$$

uz koju kriterij kvalitete postaje euklidska, odnosno L_2-norma:

$$\boxed{J(\boldsymbol{\Theta}) = \frac{1}{2}\sum_{r=1}^{N}\boldsymbol{\varepsilon}^T(r,\boldsymbol{\Theta})\boldsymbol{\varepsilon}(r,\boldsymbol{\Theta}) = \frac{1}{2}\sum_{r=1}^{N}\sum_{i=1}^{n}\varepsilon_i^2(r,\boldsymbol{\Theta}) = \frac{1}{2}\boldsymbol{\varepsilon}^{*T}(\boldsymbol{\Theta})\boldsymbol{\varepsilon}^*(\boldsymbol{\Theta})}$$

Identifikacija procesa

- **Dobra svojstva ovog kriterija kvalitete su**:
  - dvostruka derivabilnost po parametrima modela.
  - jednostavnost analize dobivenih rezultata.
- Optimalne vrijednosti parametara modela procesa $\boldsymbol{\Theta}^*$ mogu se definirati kao argument koji minimizira kriterij kvalitete:

$$\boxed{\boldsymbol{\Theta}^* = \arg\min(J(\boldsymbol{\Theta}))}$$

- Postupci vrednovanja modela procesa (model validation) predstavljaju završnu fazu postupka identifikacije procesa.
- Njihova zadaća je objektivno vrednovati identificirani model procesa, odnosno ocijeniti stupanj podudarnosti njegova ponašanja s ponašanjem stvarnoga procesa.

Identifikacija procesa

- **Usporedbu ponašanja modela procesa i stvarnoga procesa treba provoditi na podacima koji nisu korišteni za estimaciju parametara modela** (podaci za vrednovanje).
- Za vrednovanje modela procesa koriste se parametarski i korelacijski postupci.
- Parametarski postupci vrednuju identificirani model procesa uspoređujući ga s modelom veće dimenzije, pri tome se model veće dimenzije ne identificira, već se postupci vrednovanja zasnivaju na procjeni iznosa kriterija kvalitete modela veće dimenzije na temelju iznosa kriterija kvalitete identificiranoga modela.
- Korelacijski postupci vrednovanja modela temelje se na izračunavanju autokorelacijske funkcije predikcijske greške i međukorelacijskih funkcija određenih kombinacija raspoloživih signala procesa.

Estimacija parametara – model sistema

- Opći model diskretnog linearnog SISO modela s konstantnim parametrima:

$$\boxed{A(q)y(k) = \frac{B(q)}{F(q)}u(k) + \frac{C(q)}{D(q)}e(k)}$$

q^{-1} – operator kašnjenja za jedan korak diskretizacije:

$$q^{-1} f(k) = f(k-1)$$

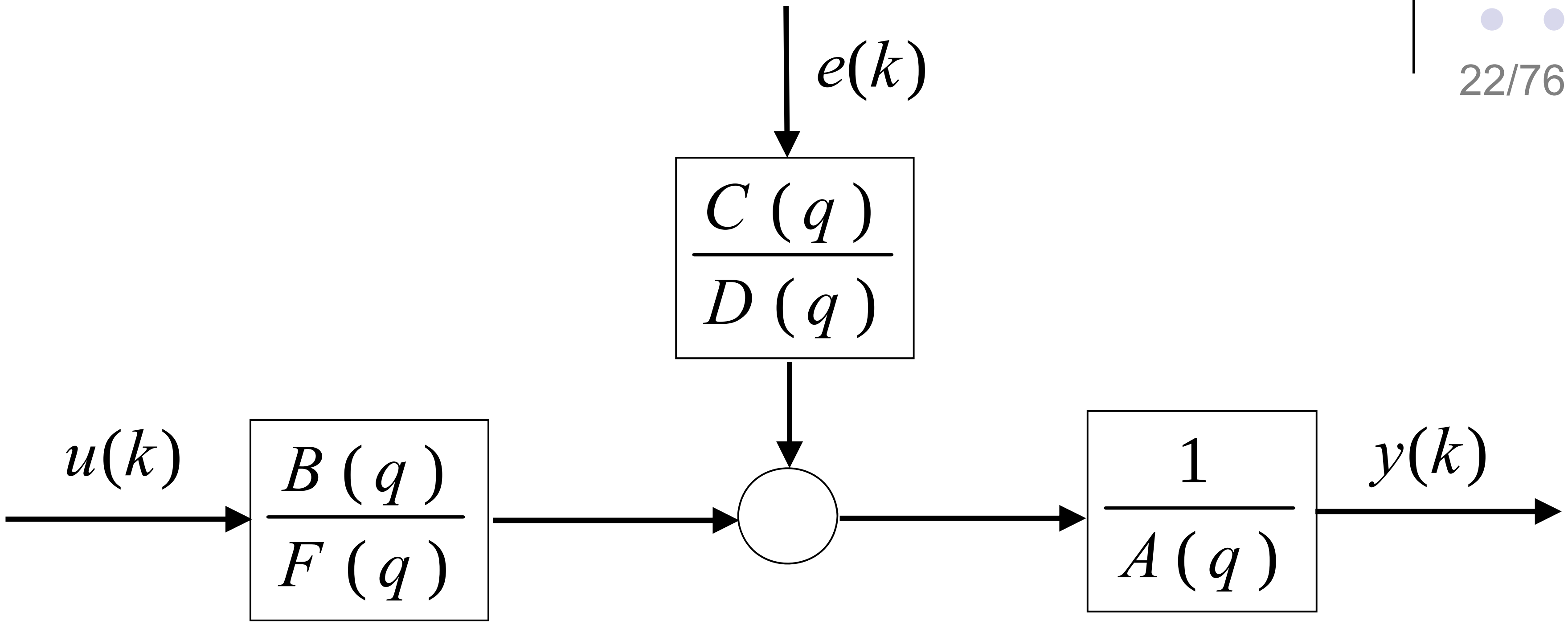
$u(k)$ – upravljački ulaz (ulazna varijabla);

$y(k)$ – izlaz procesa (izlazna varijabla);

$e(k)$ – slučajni poremećaj (nezavisna nekorelirana slučajna poremećajna varijabla sa srednjom vrijednošću 0 i standardnom devijacijom σ);

Estimacija parametara - model sistema

- Opći model diskretnog linearnog sistema:



$$A(q) = 1 + a_1 q^{-1} + \cdots + a_{na} q^{-na}, \quad B(q) = 1 + b_1 q^{-1} + \cdots + b_{nb} q^{-nb}$$

$$C(q) = 1 + c_1 q^{-1} + \cdots + c_{nc} q^{-nc}, \quad D(q) = 1 + d_1 q^{-1} + \cdots + d_{nd} q^{-nd}$$

$$F(q) = 1 + f_1 q^{-1} + \cdots + f_{nf} q^{-nf}$$

Estimacija parametara - model sistema

- Prikaz modela procesa u ARMAX obliku:

$$\boxed{A(q)y(k) = q^{-d}B(q)u(k) + C(q)e(k) + H(q)v(k) + w(k)}$$

gdje su:

$$A(q) = q^{na} + a_1 q^{na-1} + \cdots + a_{na} \qquad (6)$$

$$B(q) = b_0 q^{nb} + b_1 q^{nb-1} + \cdots + b_{nb}$$

$$C(q) = q^{nc} + c_1 q^{nc-1} + \cdots + c_{nc}$$

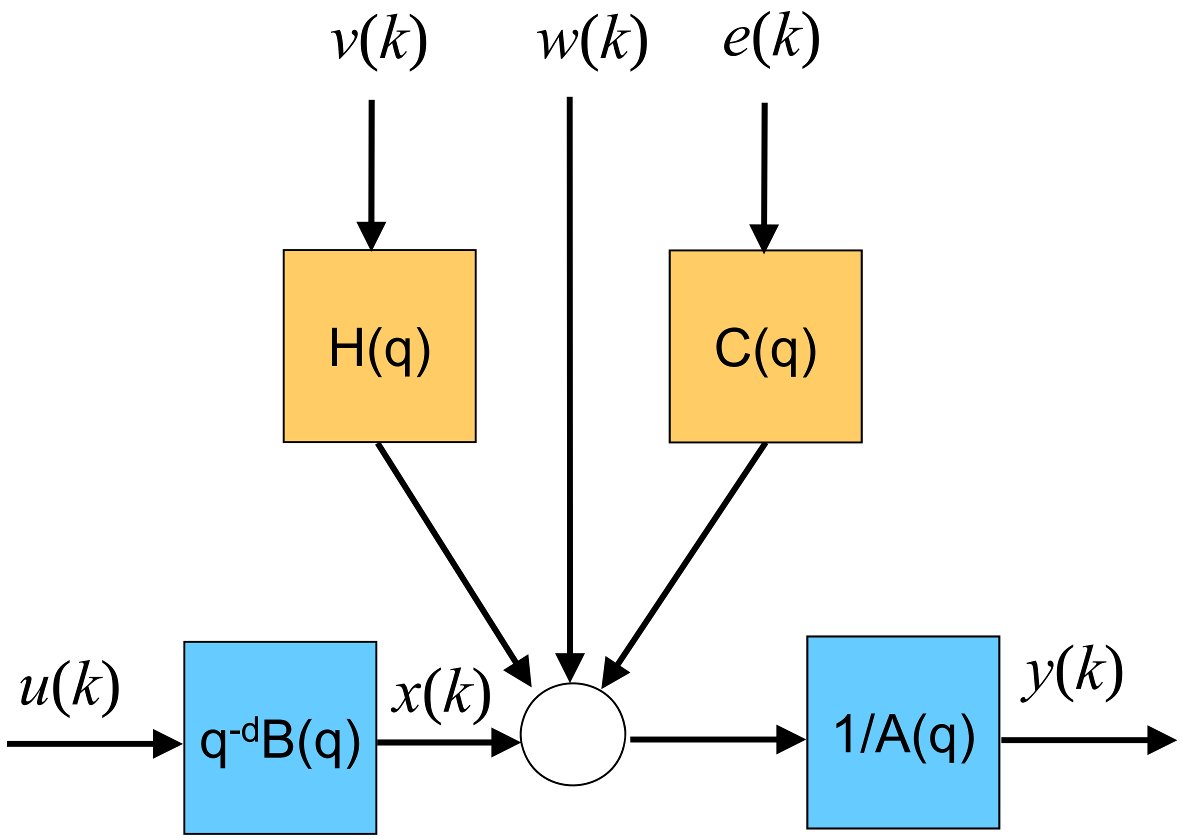
$$H(q) = h_0 q^{nh} + h_1 q^{nh-1} + \cdots + h_{nh}$$

$$w(k) = w_0 + w_1 q + w_2 q^2 + \cdots + w_{nw} q^{nw}$$

$u(k)$ - upravljački ulaz (ulazna varijabla);
$y(k)$ - izlaz procesa (izlazna varijabla);
$e(k)$ - slučajni poremećaj;
$v(k)$ - mjerljivi poremećaj; $w(k)$ - drift (posmak)

Estimacija parametara - model sistema

- Prikaz modela procesa u ARMAX obliku:



- Drift $w(k)$ je modeliran kao polinomska vremenska funkcija.
- Većina slučajeva drifta može se pokazati s $w(k) = w_0$.

Estimacija parametara - model sistema

- Alternativni prikaz modela procesa:

$$A^*(q^{-1})y(k) =$$
$$q^{-d}B^*(q^{-1})u(k) + C^*(q^{-1})e(k) + H^*(q^{-1})v(k) + w(k)$$

gdje su:

$$A^*(q^{-1}) = 1 + a_1 q^{-1} + \cdots + a_{na} q^{-na}$$
$$B^*(q^{-1}) = b_0 + b_1 q^{-1} + \cdots + b_{nb} q^{-nb}, \; b_0 \neq 0$$
$$C^*(q^{-1}) = 1 + c_1 q^{-1} + \cdots + c_{nc} q^{-nc}$$
$$H^*(q^{-1}) = h_0 + h_1 q^{-1} + \cdots + h_{nh} q^{-nh}, \; h_0 \neq 0$$

Estimacija parametara - LSM

- Parametrizacijom ARMAX modela procesa (6) u regresijskom obliku dobiva se:

$$\boxed{y(k)=\boldsymbol{\Theta}^T(k-1)\boldsymbol{\varphi}(k)+e(k)=\boldsymbol{\varphi}^T(k)\boldsymbol{\Theta}(k-1)+e(k)} \quad (7)$$

gdje su: $\boldsymbol{\Theta}$ - vektor parametara procesa, $\boldsymbol{\varphi}$ - vektor regresije, e – poremećajni signal (bijeli šum sa $E\{e\}=0$)

- **Vektor regresije**:

$$\boldsymbol{\varphi}^T(k)=[\underbrace{y_{k-1} \quad y_{k-2} \quad \cdots \quad y_{k-na}}_{na}\underbrace{u_{k-d} \quad u_{k-d-1} \quad \cdots \quad u_{k-d-nb}}_{nb+1}$$

$$\underbrace{v_k \quad v_{k-1} \quad \cdots \quad v_{k-nh}}_{nh+1}\underbrace{1 \quad k \quad \cdots \quad k^{nw}}_{nw+1}\underbrace{e_{k-1} \quad e_{k-2} \quad \cdots \quad e_{k-nc}}_{nc}]$$

Estimacija parametara - LSM

- **Vektor parametara procesa**:

$$\boldsymbol{\Theta}^T(k) = [\underbrace{-a_1 \quad -a_2 \quad \cdots \quad -a_{na}}_{na} \underbrace{b_0 \quad b_1 \quad \cdots \quad b_{nb}}_{nb+1}$$

$$\underbrace{h_0 \quad h_1 \quad \cdots \quad h_{nh}}_{nh+1} \underbrace{w_0 \quad w_1 \quad \cdots \quad w_{nw}}_{nw+1} \underbrace{c_1 \quad c_2 \quad \cdots \quad c_{nc}}_{nc}]$$

- U navedenom modelu, nepoznati parametri procesa su linearno ovisni, tj. izlazni signal linearno ovisi o parametrima.
- Ukoliko ne poznajemo $e(k)$, što predstavlja sekvencu nekoreliranih bijelih šumova, **predikcija** od $y(k)$ je:

$$\boxed{\hat{y}(k|\boldsymbol{\Theta}) = \boldsymbol{\Theta}^T(k-1)\boldsymbol{\varphi}(k)} \qquad (8)$$

Estimacija parametara - LSM

- Vektor predikcije izlazne varijable estimira njenu vrijednost uz vektor parametara Θ iz prošlih koraka: $y(k\text{-}1)$, $u(k\text{-}1)$, $i=1,2,\ldots$
- Predikcija linearno ovisi o vektoru Θ.
- Ako je $e(k)$ sekvenca neovisnih slučajnih varijabli sa $E\{e\}=0$, tj. sekvenca bijelog šuma, tada je pogreška predikcije:

$$\boxed{\varepsilon(k) = y(k) - \hat{y}(k|\boldsymbol{\Theta})}$$

- **Uz poznatu strukturu modela procesa i mjerne podatke iz vektora φ izlazna varijabla sistema opisuje se sa**:

$$\boxed{y(k) = \boldsymbol{\varphi}^T(k)\hat{\boldsymbol{\Theta}} + \hat{\varepsilon}(k)} \qquad (9)$$

Estimacija parametara - LSM

- U jednadžbi (9) oznake imaju značenja:

  $\hat{\boldsymbol{\Theta}}$ - **vektor estimiranih parametara modela**;

  $\hat{\varepsilon}(k)$ - **pogreška estimacije u trenutku k**.

- **Cilj je odrediti $\hat{\boldsymbol{\Theta}}$ takav da pogreška estimacije** $\hat{\varepsilon}(k)$ **bude minimalna**:

$$\boxed{\hat{\varepsilon}(k) = e(k) + \boldsymbol{\varphi}^T(k)[\boldsymbol{\Theta} - \hat{\boldsymbol{\Theta}}]}$$

- Ako se postigne $\boldsymbol{\Theta} = \hat{\boldsymbol{\Theta}}$ tada vrijedi:

$$\boxed{\hat{\varepsilon}(k) \approx e(k) \quad \text{za} \quad \boldsymbol{\Theta} \approx \hat{\boldsymbol{\Theta}}}$$

Estimacija parametara - LSM

- Model (9) može se zapisati u vektorskom (matričnom) obliku:

$$\begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\varphi}^T(1) \\ \boldsymbol{\varphi}^T(2) \\ \vdots \\ \boldsymbol{\varphi}^T(N) \end{bmatrix} \hat{\boldsymbol{\Theta}} + \begin{bmatrix} \hat{\varepsilon}(1) \\ \hat{\varepsilon}(2) \\ \vdots \\ \hat{\varepsilon}(N) \end{bmatrix} \tag{10}$$

- Da bi se mogao naći $\hat{\boldsymbol{\Theta}}$ mora biti $N \geq \boldsymbol{n}_{par}$ (**problem aproksimacije**), gdje je n_{par} broj estimiranih parametara u vektoru $\hat{\boldsymbol{\Theta}}$:

$$n_{par} = na + nb + 1 + nh + 1 + nw + nc$$

Estimacija parametara - LSM

- Kada je $N = n_{par}$ imamo **interpolacijski** problem i jednadžba (10) može se predstaviti kao skup linearnih jednadžbi sa n_{par} nepoznatih parametara.
- Za uspješnu estimaciju, sistem mora raditi dovoljno dugo da se formira N vektora podataka, gdje je $N \geq n_{par}$ (ili $N >> n_{par}$).
- **Linearna metoda najmanjih kvadrata (LS metoda - LSM) je najviše korištena estimacijska tehnika za ove probleme**.
- Budući da imamo više mjerenja (N) nego nepoznatih parametara (n_{par}), moramo izabrati **estimator** od Θ koji **minimizira efekt pogreški**.

Estimacija parametara - LSM

- Jednadžba (10) može se ponovo napisati u obliku (staced form):

$$\boldsymbol{Y} = \boldsymbol{\Phi}\hat{\boldsymbol{\Theta}} + \hat{\boldsymbol{E}} \qquad (11)$$

gdje su:

$$\boldsymbol{Y} = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix}; \quad \boldsymbol{\Phi} = \begin{bmatrix} \varphi^T(1) \\ \varphi^T(2) \\ \vdots \\ \varphi^T(N) \end{bmatrix}; \quad \hat{\boldsymbol{E}} = \begin{bmatrix} \hat{\varepsilon}(1) \\ \hat{\varepsilon}(2) \\ \vdots \\ \hat{\varepsilon}(N) \end{bmatrix}$$

odnosno:

$$\hat{\boldsymbol{E}} = \boldsymbol{Y} - \boldsymbol{\Phi}\hat{\boldsymbol{\Theta}} \qquad (12)$$

Estimacija parametara - LSM

- LS princip glasi:

  Suma kvadrata razlike mjernog vektora *Y* i estimiranog izlaznog vektora mora biti minimalna.

- Suma kvadrata pogreški sa jediničnim težinama (LS kriterij) može se definirati kao:

$$J(\hat{\boldsymbol{\Theta}}) = \frac{1}{2}\sum_{i=1}^{N} \hat{\varepsilon}_i^2(k) = \frac{1}{2}\hat{\boldsymbol{E}}^T \hat{\boldsymbol{E}} = \frac{1}{2}\left\|\hat{\boldsymbol{E}}\right\|^2$$

- Otežana suma kvadrata navedene razlike:

$$J(\hat{\boldsymbol{\Theta}}) = \frac{1}{2}\sum_{i=1}^{N} q_i \hat{\varepsilon}_i^2(k) = \frac{1}{2}\hat{\boldsymbol{E}}^T \boldsymbol{Q}\hat{\boldsymbol{E}}$$

Estimacija parametara - LSM

- LS kriterij, odnosno kriterij najmanjih kvadrata, je:

$$J(\hat{\boldsymbol{\Theta}}) = \frac{1}{2}\hat{\boldsymbol{E}}^T\hat{\boldsymbol{E}} = (\boldsymbol{Y} - \boldsymbol{\Phi}\hat{\boldsymbol{\Theta}})^T(\boldsymbol{Y} - \boldsymbol{\Phi}\hat{\boldsymbol{\Theta}})$$

$$J(\hat{\boldsymbol{\Theta}}) = \boldsymbol{Y}^T\boldsymbol{Y} - \boldsymbol{Y}^T\boldsymbol{\Phi}\hat{\boldsymbol{\Theta}} - \hat{\boldsymbol{\Theta}}^T\boldsymbol{\Phi}^T\boldsymbol{Y} + \hat{\boldsymbol{\Theta}}^T\boldsymbol{\Phi}^T\boldsymbol{\Phi}\hat{\boldsymbol{\Theta}}$$

- Postavljanjem na nulu derivacije od $J(\hat{\boldsymbol{\Theta}})$ s obzirom na $\hat{\boldsymbol{\Theta}}$ za stacionarnu tačku (matrica Jacobiana) dobiva se **ekstrem**:

$$\frac{\partial J(\hat{\boldsymbol{\Theta}})}{\partial \hat{\boldsymbol{\Theta}}} = \boldsymbol{\Phi}^T\boldsymbol{\Phi}\hat{\boldsymbol{\Theta}} - \boldsymbol{Y}^T\boldsymbol{\Phi} = 0 \qquad (13)$$

Estimacija parametara - LSM

- Minimum se postiže uz pozitivnu drugu derivaciju kriterija po parametrima $\hat{\boldsymbol{\Theta}}$ (Hessian):

$$\boxed{\left[\frac{\partial^2 J(\hat{\boldsymbol{\Theta}})}{\partial^2 \hat{\boldsymbol{\Theta}}}\right] = \boldsymbol{\Phi}^T \boldsymbol{\Phi} \geq 0}$$

- Hessian od $J(\hat{\boldsymbol{\Theta}})$ će biti pozitivno definitna matrica ako i samo ako $\boldsymbol{\Phi}$ ima puni rang.
- LS estimacija parametara daje (iz jednadžbe (13)):

$$\boxed{\hat{\boldsymbol{\Theta}}_{LS} = (\boldsymbol{\Phi}^T \boldsymbol{\Phi})^{-1} \boldsymbol{\Phi}^T \boldsymbol{Y} = \boldsymbol{\Phi}^{\Diamond} \boldsymbol{Y}} \qquad (14)$$

gdje je $\boldsymbol{\Phi}^{\Diamond}$ pseudoinverzija od $\boldsymbol{\Phi}$.

Estimacija parametara - LSM

- **Pseudoinverzna matrice $\boldsymbol{\Phi}$ postoji ako i samo ako matrica $\boldsymbol{\Phi}$ ima puni rang, što je zadovoljeno uz stalnu pobudu sistema test signalom**.
- Drugim riječima, pseudoinverzija od $\boldsymbol{\Phi}$ postoji ako je $\boldsymbol{\Phi}^T\boldsymbol{\Phi}$ nesingularna matrica i to će biti zadovoljeno kada $\boldsymbol{\Phi}$ ima puni rang, tj. kada je signal perzistentan (stalno pobuđujući).
- **Rezultantna pogreška ugađanja** (fitovanja), odnosno **pogreška estimacije** je:

$$\boxed{\hat{\boldsymbol{E}} = \boldsymbol{R}^T = [\eta(1)\ \ \eta(2)\ldots\eta(N)]}$$

gdje se komponente od $\boldsymbol{R}$ nazivaju **reziduima** (ostacima).

Estimacija parametara - LSM

- Iz jednadžbe $\boldsymbol{Y} = \boldsymbol{\Phi}\hat{\boldsymbol{\Theta}} + \hat{\boldsymbol{E}}$ slijedi:

$$\boxed{\hat{\boldsymbol{E}} = \boldsymbol{R}^T = \boldsymbol{Y} - \boldsymbol{\Phi}\hat{\boldsymbol{\Theta}}_{LS}}$$

- Iz prethodne jednadžbe proizlazi:

$$\boxed{\boldsymbol{\Phi}^T\boldsymbol{Y} = \boldsymbol{\Phi}^T\boldsymbol{\Phi}\hat{\boldsymbol{\Theta}}_{LS} + \boldsymbol{\Phi}^T\boldsymbol{R}}$$

što kombinirano sa:

$$\hat{\boldsymbol{\Theta}}_{LS} = (\boldsymbol{\Phi}^T\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^T\boldsymbol{Y}$$

daje:

$$\boxed{\boldsymbol{\Phi}^T\boldsymbol{R} = \boldsymbol{0}}$$

Estimacija parametara - LSM

- Iz definicije rezidua:

$$\boldsymbol{R} = (\boldsymbol{Y} - \boldsymbol{\Phi}\hat{\boldsymbol{\Theta}}_{LS})$$

i uvjeta:

$$\boldsymbol{\Phi}^T \boldsymbol{R} = \mathbf{0}$$

$$[\boldsymbol{\varphi}(1) \ \ \boldsymbol{\varphi}(2) \ \ \ldots \ \ \boldsymbol{\varphi}(N)]\boldsymbol{R} = \mathbf{0}$$

proizlazi:

$$\sum_{k=1}^{N} y(k-i)\eta(k) = 0 \quad \forall i = 1,2,\ldots,na$$

$$\sum_{k=1}^{N} u(k-i)\eta(k) = 0 \quad \forall i = 1,2,\ldots,nb+1$$

Estimacija parametara - LSM

- Za veliki N, uvjeti poprimaju oblik:

$$\boxed{\begin{aligned} E\{y(k-i)\eta(k)\} &= 0 \quad \forall i = 1,2,\dots,na \\ E\{u(k-i)\eta(k)\} &= 0 \quad \forall i = 1,2,\dots,nb+1 \end{aligned}}$$

- Za $\hat{\boldsymbol{\Theta}}_{LS}$ uzimajući podatke od 1 do k estimirani parametri imaju oblik:

$$\boxed{\hat{\boldsymbol{\Theta}}_{LS} = (\boldsymbol{\Phi}^T(k)\boldsymbol{\Phi}(k))^{-1}\boldsymbol{\Phi}^T(k)\boldsymbol{Y}(k)} \qquad (15)$$

- U trenutku $k+1$ dobivaju se nova mjerenja iz procesa:

$$\boldsymbol{Y}(k+1) = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(k) \\ y(k+1) \end{bmatrix} = \begin{bmatrix} \boldsymbol{Y}(k) \\ y(k+1) \end{bmatrix} \quad \boldsymbol{\Phi}(k+1) = \begin{bmatrix} \boldsymbol{\varphi}^T(1) \\ \boldsymbol{\varphi}^T(2) \\ \vdots \\ \boldsymbol{\varphi}^T(k) \\ \boldsymbol{\varphi}^T(k+1) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Phi}(k) \\ \boldsymbol{\varphi}^T(k+1) \end{bmatrix}$$

Estimacija parametara - LSM

- U trenutku $k+1$ estimirani parametri su:

$$\boxed{\hat{\boldsymbol{\Theta}}_{LS}(k+1) = (\boldsymbol{\Phi}^T(k+1)\boldsymbol{\Phi}(k+1))^{-1}\boldsymbol{\Phi}^T(k+1)\boldsymbol{Y}(k+1)}$$

- Slijedi da je:

$$\begin{aligned} \boldsymbol{\Phi}^T(k+1)\boldsymbol{\Phi}(k+1) &= [\boldsymbol{\Phi}^T(k) \quad \boldsymbol{\varphi}(k+1)]\begin{bmatrix} \boldsymbol{\Phi}(k) \\ \boldsymbol{\varphi}^T(k+1) \end{bmatrix} \\ &= \boldsymbol{\Phi}^T(k)\boldsymbol{\Phi}(k) + \boldsymbol{\varphi}(k+1)\boldsymbol{\varphi}^T(k+1) \end{aligned} \qquad (16)$$

- Korištenjem prethodnog izraza određuju se nove vrijednosti $\boldsymbol{\Phi}^T(k+1)\boldsymbol{\Phi}(k+1)$.
- Problem je odrediti **inverznu matricu** direktno (rekurzivnom metodom) bez potrebe za traženjem kompletne inverzne matrice u svakom koraku.
- Osim toga potrebno je izračunati i $\boldsymbol{\Phi}^T(k+1)\boldsymbol{Y}(k+1)$.

Estimacija parametara - LSM

Svojstva LS estimatora

- $\hat{\boldsymbol{\Theta}}_{LS}$ je slučajna varijabla čija se svojstva mogu analizirati jednadžbom:

$$\boxed{y(k) = \boldsymbol{\varphi}^T(k)\hat{\boldsymbol{\Theta}} + \hat{\varepsilon}(k)} \tag{17}$$

  koja definira stvarni sistem sa poremećajima.

- $\hat{\boldsymbol{\Theta}}_{LS}$ karakteriziraju:
  - **bias** (sistemska pogreška estimiranih parametara)
  - **kovarijanca** (raspršenje estimiranih parametara uzrokovanih slučajnom pogreškom).

Estimacija parametara - LSM

Bias

- Korištenjem jednadžbe (17) za $k = 1,2,\ldots, N$ dobiva se:

$$Y = \boldsymbol{\Phi}\hat{\boldsymbol{\Theta}} + \hat{\boldsymbol{E}}$$

- Uvrštavanjem u jednadžbu estimiranih parametara:

$$\hat{\boldsymbol{\Theta}}_{LS} = (\boldsymbol{\Phi}^T\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^T Y$$

dobiva se:

$$\boxed{\hat{\boldsymbol{\Theta}}_{LS} = (\boldsymbol{\Phi}^T\boldsymbol{\Phi})^{-1}[\boldsymbol{\Phi}^T\boldsymbol{\Phi}\boldsymbol{\Theta} + \boldsymbol{\Phi}^T\hat{\boldsymbol{E}}] = \boldsymbol{\Theta}(\boldsymbol{\Phi}^T\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^T\hat{\boldsymbol{E}}}$$

- Srednja devijacija estimiranih parametara od njihove stvarne vrijednosti određena je sa:

$$\boxed{\hat{\boldsymbol{\Theta}}_{LS} - \boldsymbol{\Theta} = (\boldsymbol{\Phi}^T\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^T\hat{\boldsymbol{E}}}$$

Estimacija parametara - LSM

- Kada su podaci u $\boldsymbol{\Phi}$ deterministički i srednja vrijednost $\boldsymbol{e}(k)$ jednaka nuli, očekivanje pogreške parametara je:

$$E\{\hat{\boldsymbol{\Theta}}_{LS}-\boldsymbol{\Theta}\}=(\boldsymbol{\Phi}^T\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^T E_E\{\hat{\boldsymbol{E}}\}=0$$

- Ukoliko su elementi matrice $\boldsymbol{\Phi}$ slučajni, ali nezavisni, tada imamo:

$$E\{\hat{\boldsymbol{\Theta}}_{LS}-\boldsymbol{\Theta}\}=E_{\Phi}\{(\boldsymbol{\Phi}^T\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^T\}E_E\{\hat{\boldsymbol{E}}\}=0$$

- Indeksi E i Φ označavaju da se radi o očekivanju s obzirom na $\boldsymbol{E}$ i $\boldsymbol{\Phi}$ objekte.
- **Zaključak**: neophodna je asimptotska nekoreliranost šuma e i podataka $\boldsymbol{\Phi}$ za estimaciju bez pogreške.

RLS estimator

- U nastavku se opisuje estimator zasnovan na metodi najmanjih kvadrata, tzv. **RLS** (Recursive Least Square) estimator.
- Osnovna svojstva RLS estimatora su:
  - Ne osvježavaju se svi podaci u svakom koraku k.
  - Ne računa se cijeli:

$$\boxed{\hat{\boldsymbol{\Theta}}_{LS}(k)}$$

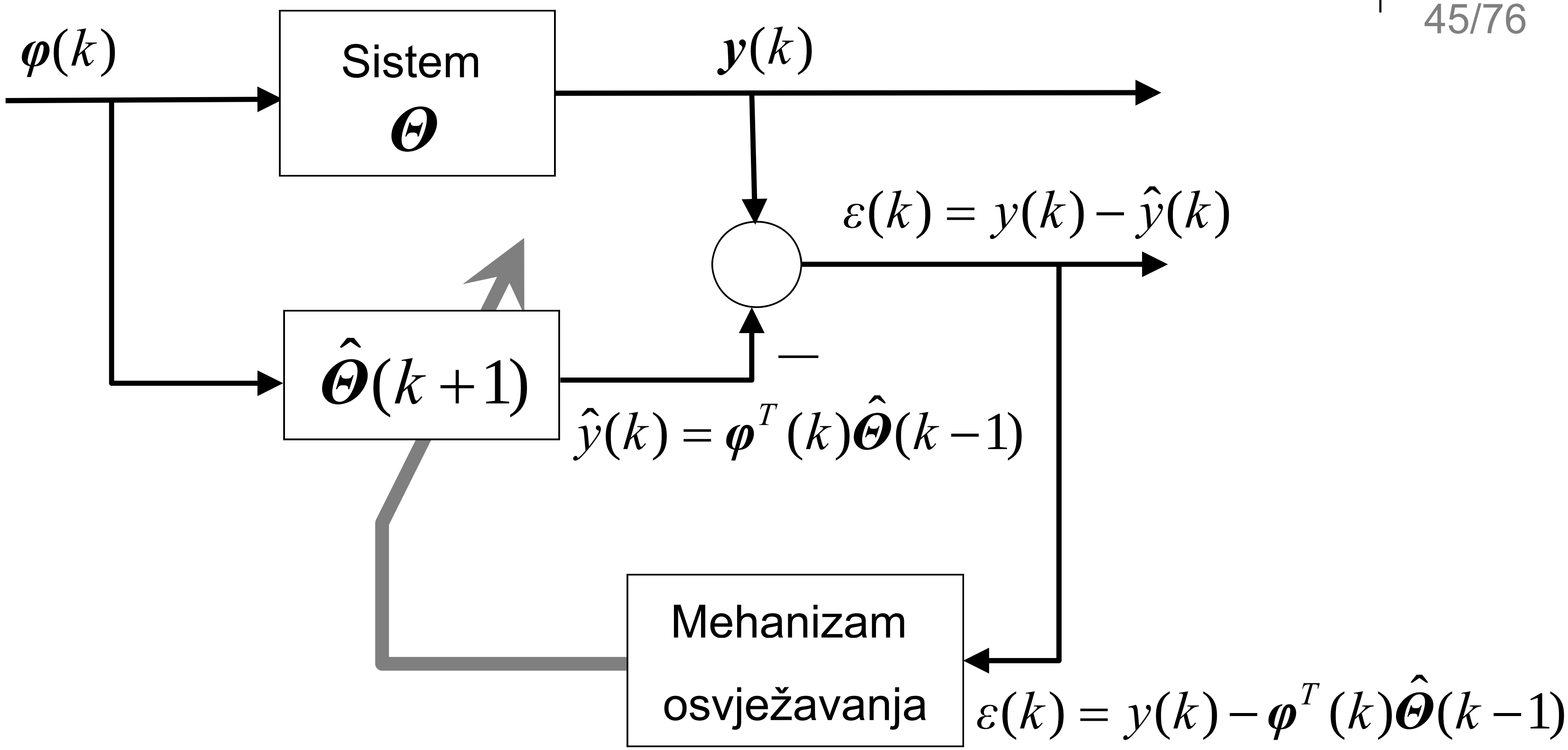
  - Dodaje se samo jedan novi podatak i računa se:

$$\boxed{\hat{\boldsymbol{\Theta}}_{LS}(k+1)}$$

  - Radi se o rekurzivnom algoritmu.

RLS estimator

- Blok dijagram RLS estimatora.



RLS estimator

- Uvrštavanjem vrijednosti za:

$$\boldsymbol{\Phi}^T(k+1) \text{ i } \boldsymbol{Y}(k+1)$$

u izraz (16) dobiva se:

$$\boldsymbol{\Phi}^T(k+1)\boldsymbol{Y}(k+1) = [\boldsymbol{\Phi}^T(k) \quad \boldsymbol{\varphi}(k+1)]\begin{bmatrix} \boldsymbol{Y}(k) \\ y(k+1) \end{bmatrix}$$

$$= \boldsymbol{\Phi}^T(k)\boldsymbol{Y}(k) + \boldsymbol{\varphi}(k+1)y(k+1)$$

- Uvođenjem oznaka: $\boldsymbol{P}(k) = [\boldsymbol{\Phi}^T(k)\boldsymbol{\Phi}(k)]^{-1}$; $\boldsymbol{B}(k) = \boldsymbol{\Phi}^T(k)\boldsymbol{Y}(k)$ jednadžbe estimiranih parametara u k i k+1 koraku postaju:

$$\hat{\boldsymbol{\Theta}}_{LS}(k+1) = \boldsymbol{P}(k+1)\boldsymbol{B}(k+1)$$

$$\hat{\boldsymbol{\Theta}}_{LS}(k) = \boldsymbol{P}(k)\boldsymbol{B}(k)$$

RLS estimator

- Za

$$\boldsymbol{\Phi}^T(k+1)\boldsymbol{\Phi}(k+1) = \boldsymbol{\Phi}^T(k)\boldsymbol{\Phi}(k) + \boldsymbol{\varphi}(k+1)\boldsymbol{\varphi}^T(k+1)$$

slijedi:

$$\boxed{\boldsymbol{P}^{-1}(k+1) = \boldsymbol{P}^{-1}(k) + \boldsymbol{\varphi}(k+1)\boldsymbol{\varphi}^T(k+1)}$$

- Za

$$\boldsymbol{\Phi}^T(k+1)\boldsymbol{Y}(k+1) = \boldsymbol{\Phi}^T(k)\boldsymbol{Y}(k) + \boldsymbol{\varphi}(k+1)\,y(k+1)$$

slijedi:

$$\boxed{\boldsymbol{B}(k+1) = \boldsymbol{B}(k) + \boldsymbol{\varphi}(k+1)\,y(k+1)}$$

- Druga jednadžba daje direktno osvježavanje (update) iz $\boldsymbol{B}(k)$ u $\boldsymbol{B}(k+1)$, kao i $\boldsymbol{P}(k+1)$ iz $\boldsymbol{P}(k)$ iz prve jednadžbe, odnosno one omogućuju **rekurzivno računanje novih vrijednosti**.

RLS estimator

- Kako se radi rekurzija i osvježavanje – **Lema inverzije matrica** (Matrix inversion lemma):

  Neka su:

$$A = P^{-1}(k),\ B = \varphi(k+1),\ C = I,\ D = \varphi^T(k+1)$$

  matrice kompatibilnih dimenzija tako da postoji produkt i suma oblika:

$$A + BCD = P^{-1}(k) + \varphi(k+1)\varphi^T(k+1)$$

  tada vrijedi:

$$\boxed{[A+BCD]^{-1} = A^{-1} - A^{-1}B[C^{-1} + DA^{-1}B]^{-1}DA^{-1}}$$

RLS estimator

Uzimajući vrijednosti matrice prema polinomima estimacije dobiva se:

$$[\boldsymbol{P}^{-1}(k)+\boldsymbol{\varphi}(k+1)\boldsymbol{\varphi}^T(k+1)]^{-1} = \boldsymbol{P}(k)-\frac{\boldsymbol{P}(k)\boldsymbol{\varphi}(k+1)\boldsymbol{\varphi}^T(k+1)\boldsymbol{P}(k)}{1+\boldsymbol{\varphi}^T(k+1)\boldsymbol{P}(k)\boldsymbol{\varphi}(k+1)}$$

- $\boldsymbol{P}(k)$ **je pozitivno definitna matrica** jednaka kovarijantnoj matrici od $\boldsymbol{\Theta}(k)$ s očekivanjem $\hat{\boldsymbol{\Theta}}_{LS}$.
- **Ovom lemom riješen je problem inverzije matrice u svakom koraku** (dijeljenje sa skalarom u svakom koraku).

RLS estimator

- Temeljem prethodne leme imamo:

$$\boldsymbol{P}(k+1)=\boldsymbol{P}(k)-\frac{\boldsymbol{P}(k)\boldsymbol{\varphi}(k+1)\boldsymbol{\varphi}^T(k+1)\boldsymbol{P}(k)}{1+\boldsymbol{\varphi}^T(k+1)\boldsymbol{P}(k)\boldsymbol{\varphi}(k+1)} \quad (18)$$

- Smjenom $\boldsymbol{y}(k+1)$ iz jednadžbe

$$\hat{\varepsilon}(k+1)=y(k+1)-\boldsymbol{\varphi}^T(k+1)\hat{\boldsymbol{\Theta}}_{LS}(k)=y(k+1)-\hat{y}(k+1)$$

u jednadžbu:

$$\boldsymbol{B}(k+1)=\boldsymbol{B}(k)+\boldsymbol{\varphi}(k+1)y(k+1)$$

dobiva se:

$$\boldsymbol{B}(k+1)=\boldsymbol{B}(k)+\boldsymbol{\varphi}(k+1)\boldsymbol{\varphi}^T(k+1)\hat{\boldsymbol{\Theta}}_{LS}(k)+\boldsymbol{\varphi}(k+1)\varepsilon(k+1)$$

RLS estimator

- Iz jednadžbe

$$\boldsymbol{P}^{-1}(k+1)\hat{\boldsymbol{\Theta}}_{LS}(k+1) = \left[\underbrace{\boldsymbol{P}^{-1}(k) + \boldsymbol{\varphi}(k+1)\boldsymbol{\varphi}^{T}(k+1)}_{\boldsymbol{P}^{-1}(k+1)}\right]\hat{\boldsymbol{\Theta}}_{LS}(k) + \boldsymbol{\varphi}(k+1)\varepsilon(k+1)$$

množenjem sa P(k+1) dobiva se:

$$\boxed{\hat{\boldsymbol{\Theta}}_{LS}(k+1) = \hat{\boldsymbol{\Theta}}_{LS}(k) + \underbrace{\boldsymbol{P}(k+1)\boldsymbol{\varphi}(k+1)}_{\boldsymbol{K}(k+1)}\varepsilon(k+1)} \qquad (19)$$

gdje je $\boldsymbol{K}(k+1)$ **matrica pojačanja** (faktor pojačanja).

RLS estimator

RLS estimator - sumarno

$$\hat{\boldsymbol{\Theta}}_{LS}(k+1) = \hat{\boldsymbol{\Theta}}_{LS}(k) + \boldsymbol{P}(k+1)\boldsymbol{\varphi}(k+1)\varepsilon(k+1) \tag{20}$$

$$\varepsilon(k+1) = y(k+1) - \boldsymbol{\varphi}^T(k+1)\hat{\boldsymbol{\Theta}}_{LS}(k) \tag{21}$$

$$\boldsymbol{P}(k+1) = \boldsymbol{P}(k) - \frac{\boldsymbol{P}(k)\boldsymbol{\varphi}(k+1)\boldsymbol{\varphi}^T(k+1)\boldsymbol{P}(k)}{1+\boldsymbol{\varphi}^T(k+1)\boldsymbol{P}(k)\boldsymbol{\varphi}(k+1)} \tag{22}$$

RLS estimator

RLS estimator – Pseudokod

K$_1$: Inicijalizacija algoritma sa $\boldsymbol{P}(0),\ \hat{\boldsymbol{\Theta}}(0)$

K$_2$: Formiranje $\boldsymbol{\varphi}(k+1)$ iz novih mjerenih podataka

K$_3$: Formiranje $\varepsilon(k+1)$ iz novih mjerenih podataka

$$\boxed{\varepsilon(k+1) = y(k+1) - \boldsymbol{\varphi}^T(k+1)\hat{\boldsymbol{\Theta}}(k)}$$

K$_4$: Izračunati $\boldsymbol{P}(k+1)$

$$\boxed{\boldsymbol{P}(k+1) = \boldsymbol{P}(k) - \frac{\boldsymbol{P}(k)\boldsymbol{\varphi}(k+1)\boldsymbol{\varphi}^T(k+1)\boldsymbol{P}(k)}{1+\boldsymbol{\varphi}^T(k+1)\boldsymbol{P}(k)\boldsymbol{\varphi}(k+1)}}$$

K$_5$: Izračunati $\hat{\boldsymbol{\Theta}}(k+1)$

$$\boxed{\hat{\boldsymbol{\Theta}}(k+1) = \hat{\boldsymbol{\Theta}}(k) + \boldsymbol{P}(k+1)\boldsymbol{\varphi}(k+1)\varepsilon(k+1)}$$

K$_6$: Čekati sljedeći korak diskretizacije i ići na ***K$_2$***.

RLS estimator

- Navedene jednadžbe (jednadžbe RLS estimatora) mogu se zapisati i u sljedećem obliku:

$$\begin{aligned}
\boldsymbol{K}(k+1) &= \boldsymbol{P}(k)\boldsymbol{\varphi}(k+1)[\boldsymbol{I}+\boldsymbol{\varphi}^T(k+1)\boldsymbol{P}(k)\boldsymbol{\varphi}(k+1)]^{-1} \\
\boldsymbol{P}(k+1) &= [\boldsymbol{I}-\boldsymbol{K}(k+1)\boldsymbol{\varphi}^T(k+1)]\boldsymbol{P}(k) \\
\hat{\boldsymbol{\Theta}}(k+1) &= \hat{\boldsymbol{\Theta}}(k)+\boldsymbol{K}(k+1)[y(k+1)-\boldsymbol{\varphi}^T(k+1)\hat{\boldsymbol{\Theta}}(k)]
\end{aligned} \tag{23}$$

RLS estimator

Svojstva RLS algoritma:

- Estimira samo koeficijente od $\boldsymbol{A}$ i $\boldsymbol{B}$.
- Ne daje dobre rezultate ako se koristi obojeni šum.
- Ne pohranjuju se svi podaci, mali zahtjevi na memoriju.
- Prikladan za on-line estimaciju.
- Nije potrebno, za razliku od LS-a, računati inverziju matrice $\boldsymbol{\Phi}$.
- Problem: izbor početnih vrijednosti $\hat{\boldsymbol{\Theta}}(0),\ \boldsymbol{P}(0)$.
- Jednostavno modificirati u real-time algoritam.
- Centralni dio adaptivnog sistema.
- Koristi se u detekciji kvarova kako bi se prepoznale značajnije promjene u sistemu.

RLS estimator

Željena svojstva rekurzivnih algoritama:

- Brza konvergencija.
- Konzistentne estimirane varijable (u slučaju vremenski invarijantnog sistema).
- Dobro slijeđenje (u slučaju vremenski invarijantnog sistema).
- Računarski jednostavan.

Trade-off

- Konvergencija vs. slijeđenje.
- Računarska složenost vs. tačnost.

Inicijalni uvjeti RLS estimatora

- Inicijalizacija početnih parametara estimatora:
  - **Vektor regresije** $\varphi(0)$.
  - **Početni vektor parametara** $\hat{\boldsymbol{\Theta}}_{LS}(0)$.
  - **Početna matrica kovarijance** $\boldsymbol{P}(0)$.
- Ako je $\boldsymbol{P}(0)$ malo tada će $\boldsymbol{K}(k+1)$ biti malo i $\hat{\boldsymbol{\Theta}}_{LS}(k+1)$ se neće mnogo promijeniti.
- Ako $\boldsymbol{P}(0)$ poprima velike vrijednosti, $\hat{\boldsymbol{\Theta}}_{LS}(k+1)$ će se brzo mijenjati i udaljavati od $\hat{\boldsymbol{\Theta}}_{LS}(0)$.
- U praksi se najčešće uzima:

$$\hat{\boldsymbol{\Theta}}_{LS}(0) = \mathbf{0}, \quad \boldsymbol{P}(0) = \rho \boldsymbol{I}$$

  gdje je ρ konstanta.
- Veliki iznosi ρ su dobri ako je inicijalna vrijednost od $\hat{\boldsymbol{\Theta}}_{LS}(0)$ neizvjesna.

Inicijalni uvjeti RLS estimatora

Efekt inicijalnih vrijednosti

- **Primjer 1**. Promatra se sistem:

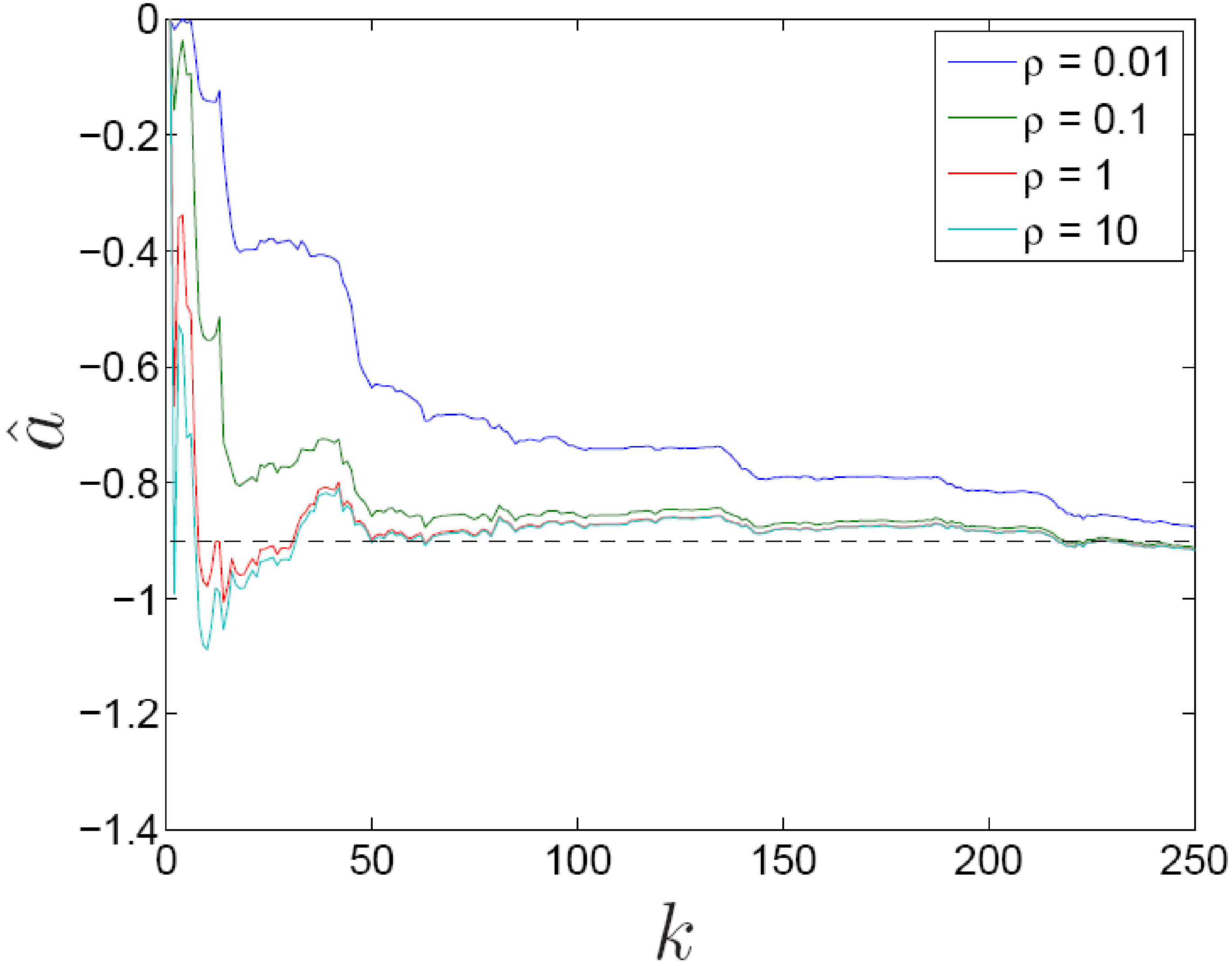
$$y(k) - 0.9y(k-1) = 1.0u(k-1) + e(k)$$

- $u(k)$ je binarni bijeli šum
- $e(k)$ je bijeli šum srednje vrijednosti 0 i varijance 1.
- Identificirati sistem korištenjem RLS metode sa 250 tačaka (podataka).
- Parametri su incijalizirani sa:

$$\hat{\Theta}(0) = 0, \quad \boldsymbol{P}(0) = \rho \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{za } \rho = 0.01,\ 0.1,\ 1,\ 10$$

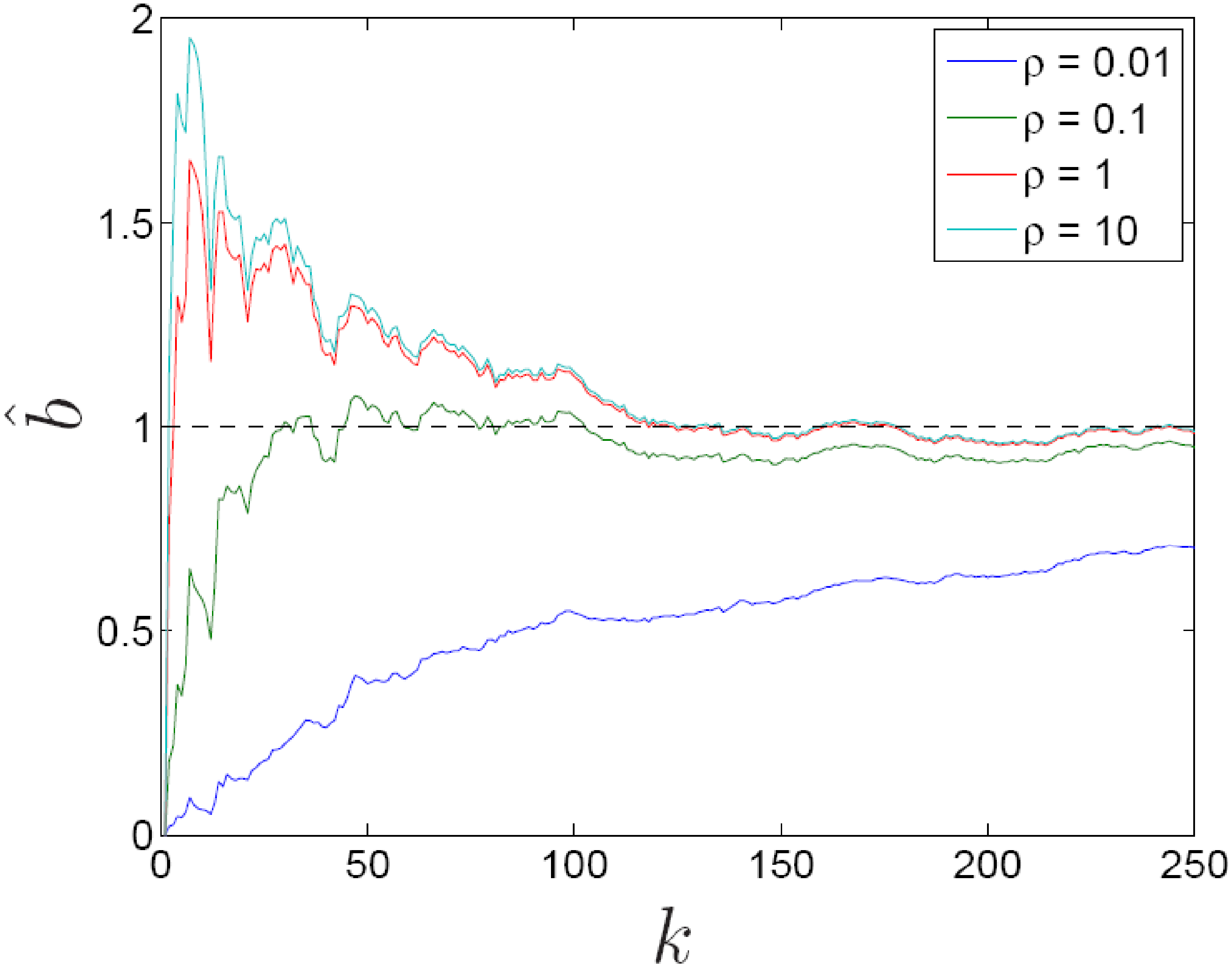
Inicijalni uvjeti RLS estimatora

Efekt inicijalnih vrijednosti



Inicijalni uvjeti RLS estimatora

Efekt inicijalnih vrijednosti



Inicijalni uvjeti RLS estimatora

- Velike i umjerene vrijednosti ρ (npr. $\rho = 1$ i $\rho = 10$) vode ka sličnim rezultatima.
- Za veliki ρ malo povjerenje je dano za $\hat{\boldsymbol{\Theta}}(0)$ tako da imamo brz tranzijentni odziv.
- Male vrijednosti ρ prouzrokuju mali $\boldsymbol{K}(k)$, što daje sporu konvergenciju.

Algoritmi za sisteme s promjenjivim param.

- **Opisani LS algoritam i iz njega izvedeni RLS algoritam dobro funkcioniraju u slučaju konstantnih, odnosno, nepromjenjivih parametara sistema**.
- Da bi se navedeni RLS algoritam moga primijeniti na sisteme sa promjenjivim parametrima mogu se načiniti sljedeće modifikacije, odnosno poboljšanja:
  - **Eksponencijalno otežavanje podataka.**
  - **Automatska promjena eksponencijalnog faktora zaboravljanja.**
  - **Resetiranje matrice kovarijance.**
  - **Modificiranje matrice kovarijance.**

RLS estimator sa faktorom zaboravljanja

- Kriterijska funkcija u LS metodi se modificira na sljedeći način:

$$J(\boldsymbol{\Theta},t)=\frac{1}{2}\sum_{k=1}^{t}\lambda^{t-k}\left(y(k)-\boldsymbol{\varphi}^{T}(k)\boldsymbol{\Theta}\right)^{2} \qquad (24)$$

  gdje je λ **faktor zaboravljanja** $(0<\lambda\le 1)$.
- Na ovaj način uvedeno je **vremenski promjenjivo otežavanje podataka**.
- Najsvježiji podaci su dani jediničnim otežavanjem, a podaci koji su stari n vremenskih jedinica sa λ^n.
- **Manje vrijednosti λ znače da će prethodne vrijednosti biti brže zaboravljene**.
- Parametri se adaptiraju tako da opisuju najnovije podatke.

RLS estimator sa faktorom zaboravljanja

- RLS metoda sa eksponencijalnim faktorom zaboravljanja postaje:

$$
\begin{aligned}
\boldsymbol{K}(k+1) &= \boldsymbol{P}(k)\boldsymbol{\varphi}(k+1)[\lambda\boldsymbol{I} + \boldsymbol{\varphi}^T(k+1)\boldsymbol{P}(k)\boldsymbol{\varphi}(k+1)]^{-1} \\
\boldsymbol{P}(k+1) &= \frac{[\boldsymbol{I} - \boldsymbol{K}(k+1)\boldsymbol{\varphi}^T(k+1)]\boldsymbol{P}(k)}{\lambda} \\
\hat{\boldsymbol{\Theta}}(k+1) &= \hat{\boldsymbol{\Theta}}(k) + \boldsymbol{K}(k+1)[y(k+1) - \boldsymbol{\varphi}^T(k+1)\hat{\boldsymbol{\Theta}}(k)]
\end{aligned}
\tag{25}
$$

- Nedostatak ove metode, RLS sa eksponencijalnim zaboravljanjem, jest da se **podaci odbacuju** čak i ako je $\boldsymbol{P}(k+1)\boldsymbol{\varphi}(k+1) = \boldsymbol{0}$.
- Ovo implicira da $y(k+1)$ ne sadrži nikakvu novu informaciju o parametru $\boldsymbol{\Theta}$.
- Iz jednadžbi slijedi da $\boldsymbol{P}$ **raste eksponencijalno** sa λ.

RLS estimator sa faktorom zaboravljanja

- Rješenje $\hat{\boldsymbol{\Theta}}(k+1)$ koje minimizira $\boldsymbol{J}(\boldsymbol{\Theta},t)$ je:

$$\hat{\boldsymbol{\Theta}}(k+1)=\left(\sum_{k=1}^{t}\lambda^{t-k}\boldsymbol{\varphi}^{T}(k+1)\boldsymbol{\varphi}(k+1)\right)^{-1}\left(\sum_{k=1}^{t}\lambda^{t-k}\boldsymbol{\varphi}^{T}(k+1)y(k+1)\right)$$

- Formula za osvježavanje slijedi analogiju sa RLS uvođenjem:

$$\boldsymbol{P}(k+1)=\left(\sum_{k=1}^{t}\lambda^{t-k}\boldsymbol{\varphi}^{T}(k+1)\boldsymbol{\varphi}(k+1)\right)^{-1}$$

- Izbor λ predstavlja trade-off između konvergencije i slijeđenja.

RLS estimator sa faktorom zaboravljanja

- Ako je λ mali $\Rightarrow$ stari podaci se brzo zaboravljaju, povlači za sobom **dobro slijeđenje**.
- Ukoliko je λ blizu 1 $\Rightarrow$ **dobra konvergencija** i male varijanse estimiranih parametara.
- U navedenom algoritmu λ se mijenja od 0.95 do 0.99.
- Imamo veću brzinu algoritma estimacije u početku.

RLS estimator sa faktorom zaboravljanja

- **Optimalan izbor λ** uzima u obzir:
  - ➢ brzo praćenje promjene parametara (veća osjetljivost na šum),
  - ➢ što manje oscilacije estimiranih parametara.
- **Preporuke za odabir λ**:
  - brze promjene parametara procesa (brzo zaboravljanje mjerenih signala),
  - spore promjene parametara ili ako nema perzistentne pobude (sporo zaboravljanje).

RLS estimator sa faktorom zaboravljanja

- **Primjer 2**. Promatra se problem slijeđenja vremenski promjenjivog sistema:

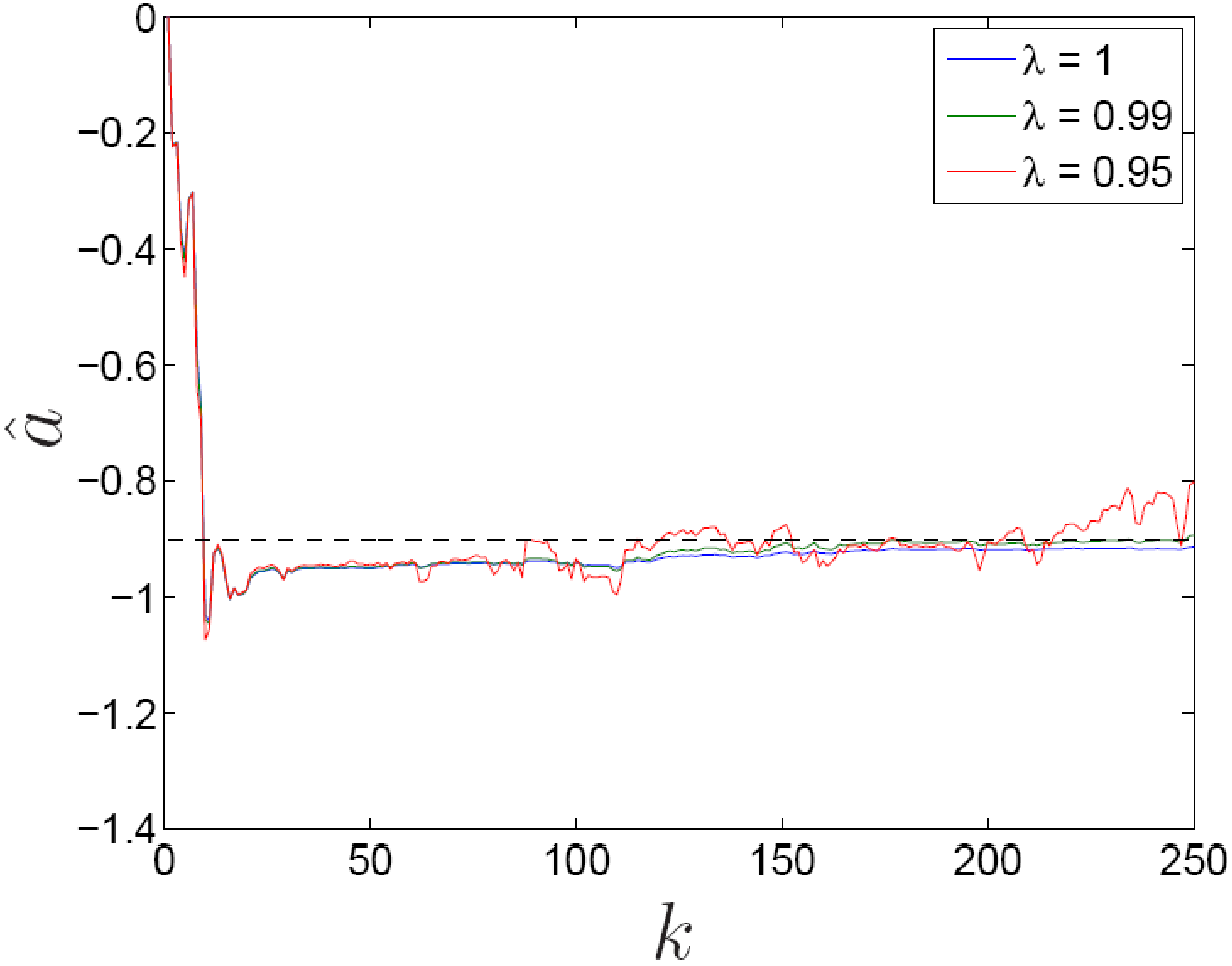
$$y(k)-0.9y(k-1)=b_0u(k-1)+e(k),\ b_0=\begin{cases}1.5 & t\le N/2\\ 0.5 & t>N/2\end{cases}$$

- $u(k)$ je binarni bijeli šum
- $e(k)$ je bijeli šum srednje vrijednosti 0 i varijance 1.
- Identificirati sistem korištenjem RLS metode sa 250 tačaka (podataka).
- Parametri su incijalizirani sa:

$$\hat{\boldsymbol{\Theta}}(0)=\mathbf{0},\quad \boldsymbol{P}(0)=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\qquad \text{za } \lambda=1,\ 0.99,\ 0.95$$

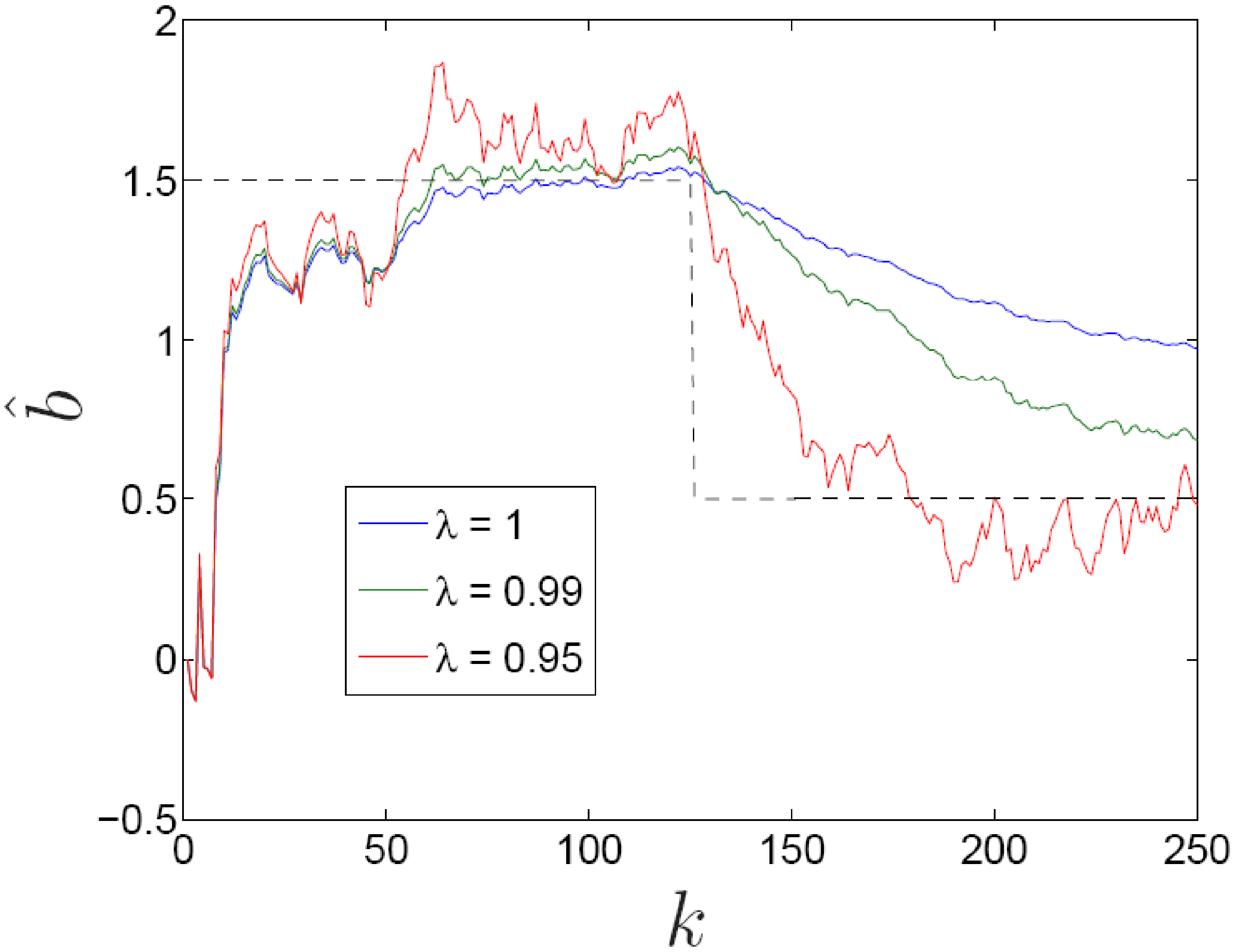
RLS estimator sa faktorom zaboravljanja

- Utjecaj promjene faktora zaboravljanja



RLS estimator sa faktorom zaboravljanja

- Utjecaj promjene faktora zaboravljanja



RLS estimator sa faktorom zaboravljanja

Analiza rezultata

- Smanjenje faktora zaboravljanja - dva efekta:
  - **dostiže tačne vrijednosti mnogo brže**,
  - **algoritam postaje mnogo osjetljiviji na šum**.
- Kako se λ smanjuje oscilacije postaju veće.
- Za postizanje konvergencije potrebno $\lambda = 1$.
- Ako je $\lambda < 1$ estimirani parametri se mijenjaju bryo I algoritam postaje mnogo osjetljiviji na šum.
- **Zbog navedenih razloga često je dobro faktor zaboravljanja mijenjati u vremenu, odnosno, učiniti ga vremenski promjenjivim**.

RLS estimator sa faktorom zaboravljanja

- **Primjer 3**. Promatra se sistem

$$y(k)-ay(k-1)=bu(k-1)+e(k)+ce(k-1)$$

gdje su:

a = -0.8, b = 0.5, $e(k)$ je bijeli šum jedinične srednje vrijednosti sa standardnom devijacijom $\sigma = 0.5$, $c = 0$ i

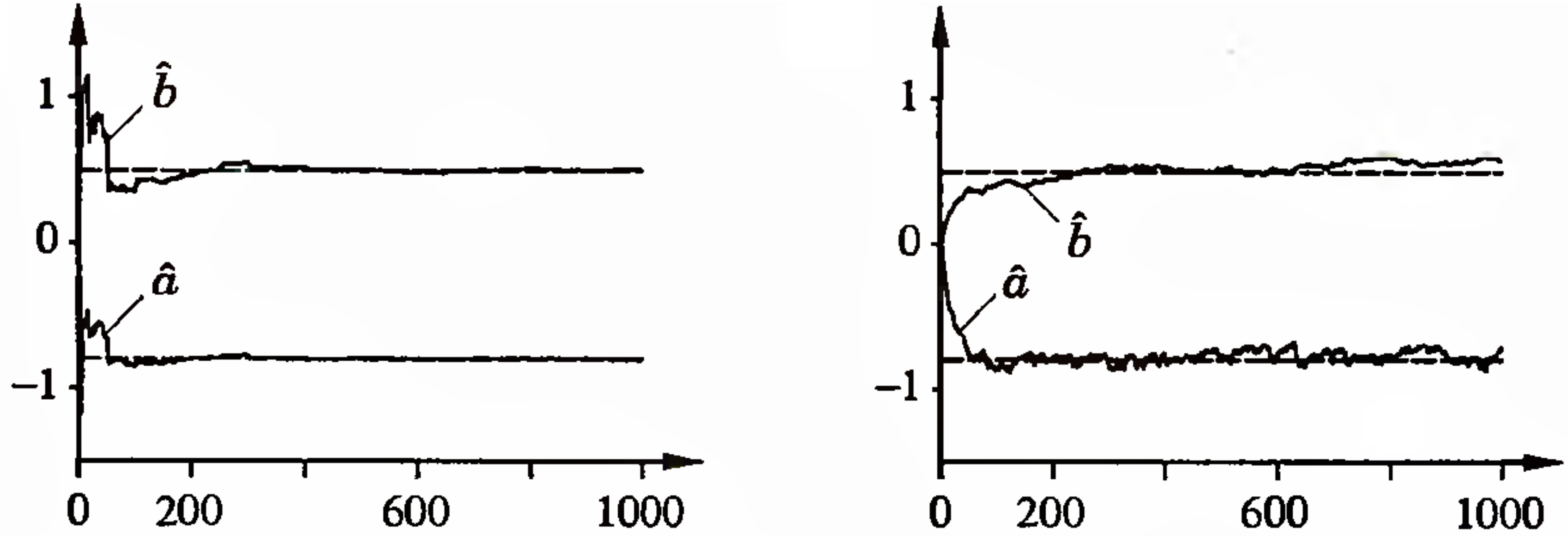
$$\hat{\boldsymbol{\Theta}}(0)=\mathbf{0},\quad \boldsymbol{P}(0)=100\cdot\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

- Vektori parametara i regresije su:

$$\hat{\boldsymbol{\Theta}}=\begin{bmatrix}\hat{a}\\ \hat{b}\end{bmatrix},\quad \boldsymbol{\varphi}(k-1)=\begin{bmatrix}-y(k-1) & u(k-1)\end{bmatrix}$$

RLS estimator sa faktorom zaboravljanja

- Rezultati dobiveni sa RLS metodom sa $\lambda = 1$ i LMS metodom sa $\gamma = 0.01$ i prikazani su respektivno na slikama a) i b).



(a) (b)

RLS estimator sa faktorom zaboravljanja

- Za vremensku promjenu parametra λ tipičan izbor je učiniti da $\lambda(k)$ teži eksponencijalno ka 1:

$$\lambda(k) = 1 - \lambda_0^k (1 - \lambda(0))$$

- Ovo se može jednostavno rekurzivno implementirati:

$$\lambda(k) = \lambda_0 \lambda(k-1) + (1 - \lambda_0), \text{ uz } \lambda_0 = 0.99, \lambda(0) = 0.95$$

- Ova vremenska promjenjivost se zatim uključi u RLS algoritam:

$$\boldsymbol{P}(k+1) = \left[\boldsymbol{I} - \frac{\boldsymbol{P}(k)\boldsymbol{\varphi}(k+1)\boldsymbol{\varphi}^T(k+1)}{\lambda(k+1) + \boldsymbol{\varphi}^T(k+1)\boldsymbol{P}(k)\boldsymbol{\varphi}(k+1)} \right] \frac{\boldsymbol{P}(k)}{\lambda(k+1)} \quad (26)$$

Eksponencijalno otežavanje podataka

- Kriterij za izvod algoritma prelazi iz:

$$J(\hat{\boldsymbol{\Theta}}) = \frac{1}{2}\sum_{i=1}^{N} q_i \hat{\varepsilon}_i^2(k) = \frac{1}{2}\hat{\boldsymbol{E}}^T \boldsymbol{Q}\hat{\boldsymbol{E}}$$

u

$$J(\hat{\boldsymbol{\Theta}}) = \frac{1}{2}\sum_{i=1}^{N} \lambda^{k-1} q_i \hat{\varepsilon}_i^2(k) \qquad (27)$$

- Odabir λ određen:
  - **brzom adaptacijom**
  - **dobrom estimacijom**.

Eksponencijalno otežavanje podataka

- Algoritam matrice kovarijance:

$$\boldsymbol{P}(k+1)=\left[\boldsymbol{I}-\frac{\boldsymbol{P}(k)\boldsymbol{\varphi}(k+1)\boldsymbol{\varphi}^T(k+1)}{\lambda(k+1)+\boldsymbol{\varphi}^T(k+1)\boldsymbol{P}(k)\boldsymbol{\varphi}(k+1)}\right]\frac{\boldsymbol{P}(k)}{\lambda(k+1)} \quad (27)$$

- $\lambda(k)$ – eksponencijalni faktor zaboravljanja:

$$0<\lambda<1$$

- Prema iskustvu najbolje je uzeti:

$$0.95<\lambda<0.99$$

- Preporučen izbor:

$$\lambda=0.98$$