

Adaptiivsed filtrid ning nende rakendused sidesüsteemides

Tõnu Trump



Toetab Euroopa Liit

SISUKORD

1. SISSEJUHATUS.....	4
2. ADAPTIIVSETE FILTRITE RAKENDUSED	5
3. WIENERI FILTRID	8
ORTOGONAALSUSE PRINTSIIP	9
WIENER-HOPFI VÖRRANDID	12
VEA PIND	14
LINEAARSE PIIRANGUGA ÜLESANNE.....	19
4. KIIREIMA LANGUSE MEETOD	25
KIIREIMA LANGUSE MEETODI RAKENDUS WIENERI FILTRILE	26
KIIREIMA LANGUSE MEETODI STABIILSUS.....	28
5. ADAPTIIVSED LMS FILTRID	35
LMS ALGORITMI ANALÜÜS	39
RUSIKAREEGLID	41
NORMALISEERITUD LMS ALGORITM	52
LMS ALGORIMI VARIANDID.....	55
6. VÄHIMRUUTUDE MEETOD	59
AJAS KESKMISTATUD KORRELATSIOONIMAATRIKSI OMADUSED	62
PROJEKTSIOONI OPERAATOR	63
VÄHIMRUUTUDE HINNANGU OMADUSED.....	63
LAHUTUS SINGULAARVÄÄRTUSTE JÄRGI	66
7. REKURSIIVNE VÄHIMRUUTUDE MEETOD	69
ANALÜÜS	75
LIBISEVA AKNAGA RLS	77
8. ALGORITMID LMS JA RLS VAHEL	78
NEWTON-RAPHSONI MEETOD	79
AFIINSE PROJEKTSIOONI ALGORITM.....	81
9. SAGEDUSVALLAS TOIMIVAD ADAPTIIVSED ALGORITMID.....	83
PIIRAMATA GRADIENDIGA ALGORITM	94
SAGEDUSVALLA ALGORITMID KOONDUVUSE KIIRENDAMISEKS.....	94
LÖPPSÕNA.....	98
INDEKS	99
KASUTATUD KIRJANDUS	100

1. Sissejuhatus

Käesolev õppevahend on pühendatud adaptiivsetele filtritele. See on ala, mis võimaldab leida lahendused mitmetele statistilise signaalitöötamise ülesannetele, nõudmata statistilise signaalitöötamise harjumuspäraseid eeldusi muutujate statistiliste omaduste nagu jaotustihedusfunktsioonid, korrelatsioonifunktsioonid jne kohta. Adaptiivsed filtrid on tänapäeval leidnud rakenduse paljude inseneriülesannete lahendamisel (telekommunikatsioon, radar, sonar jne). Samas ei ole adaptiivsete filtrite käitumise mõistmine lõpuni lihtne. On ju tegu süsteemidega, mis on ühel ajal ajas mitteinvariantsed, mittelineaarsed kui ka juhuslikud. Seetõttu oleme õppevahendis põhirõhu pannud algoritmide tuletamisele ja nende töö kirjeldamisele. Algoritmide analüüsimisel oleme esitanud vaid mõningaid üksikuid tulemusi ning selle kõrval illustreerinud algoritmide käitumist näidete varal. Näidete valikul oleme, nagu õppevahendi pealkirigi ütleb, eelistanud sidesüsteemidest pärinevaid rakendusi.

Üldjuhul me eeldame, et signaalid on komplekssignaalid, nagu on kompleksed väärtused kõrgsagedussignaali kompleksamplituudil. Selline eeldus teeb käsitluse üldisemaks. Samas piisab rakendustes, kus signaalid on reaalsed, mitte komplekssete väärtustega, taoliselt leitud algoritmide kasutamiseks signaalide imaginaarosa võtmisest võrdseks nulliga.

Adaptiivsete filtrite teooria aluseks olevad statistilise signaalitöötamise vahendid, näiteks Wieneri filter, on välja töötatud nii pidevate kui ka diskreetsete signaalide tarvis. Käesolevas õppevahendis eeldame siiski, et meie adaptiivsed filtrid toimivad diskreetses ajas. Sama kehtib tänapäeval reeglina ka adaptiivsete filtrite praktiliste rakenduste kohta.

Mõistet filter kasutame me väga üldises tähenduses sünonüümina süsteemile. Seega ei ole filter meie käsitluses mitte traditsiooniline kõrg- või madalpääsfilter, vaid süsteem, millel võib olla suvaline impulsskarakteristik. Filtri väljundsignaal on määratud konvolutsioonisummaga ning selles summas figureerivad impulsskarakteristikut iseloomustavaid kordajaid oleme tekstis vaheldumisi nimetanud kas filtri koefitsientideks või filtri kaaludeks. Käesolevas õppevahendis vaadeldud adaptiivsedid filtreid iseloomustab lõplik impulsskarakteristik.

Õppevahendi koostamisel on kasutatud mitmeid tekste, millele leiab viited kirjanduse loetelust. Suurim mõju käesolevale õppevahendile on olnud allikal [5], mida on mõnel pool kasutatud ka ilma konkreetse viiteta tekstis.

Õppevahendis kasutame läbivalt järgmisi tähistusi: skalaarsed muutujad on esitatud kaldkirjas, vektorid on tähistatud väiketähtedega paksus püstkirjas ning maatriksid suurtähtedega paksus püstkirjas. Ülemine indeks „T“ tähistab maatriksi transpositsiooni. Ülemine indeks „*“ tähistab kompleksarvu kaaskompleksi ja ülemine indeks „H“ maatriksi Hermiti transpositsiooni. Operaatoriga $E[\cdot]$ tähistame juhusliku suuruse üle ansambli keskmistatud keskvaartust ehk matemaatilist ootust, seega $E[x] = \int_{-\infty}^{\infty} xf(x)dx$, kus $f(x)$ on juhusliku suuruse x tõenäosustihedusfunktsioon. Operaator $Tr[\cdot]$ tähistab maatriksi jälge ehk maatriksi peadiagonaalil paiknevate elementide summat.

Õppevahendis esinevad joonised võib leida originaalkujul st värvilisena internetileheküljelt <http://www.lr.ttu.ee/~ttrump/adaptiivneST>.

Autor on tänulik Katre Woortile, Jürgo Predenile ja Aleksandr Tišinile käsikirja kohta tehtud asjalike märkuste eest.

Õppevahendi väljaandmist rahastati Tallinna Tehnikaülikooli projektist "Välisõppejõu kutsumine TTÜ raadio- ja sidetehnika instituuti eesmärgiga arendada telekommunikatsiooni eriala magistriõppes", projekti tegevused viiakse ellu EL struktuurifondide meetme 1.1 raames (tööjõu paindlikkust, toimetulekut ja elukestvate õpet tagav ning kõigile kättesaadav haridussüsteem). [SA Innove](#).

2. Adaptiivsete filtrite rakendused

Adaptiivsete filtrite võime kohanduda tööks tundmatus keskkonnas ning järgida selle keskkonna muutusi teeb neist käepärase vahendi mitmetest valdkondadest pärit ülesannete lahendamisel. Vaatamata sellele, et adaptiivsed filtrid leiavad kasutamist väga erinevates valdkondades, on neil siiski üks oluline ühine joon. Nimelt leiab adaptiivne filter kõikides rakendustes filtri koefitsiendid, mille rakendamisel sisendsignaali formeeritakse sisendist uus, teatud soovitud signaaliga võimalikult sarnane signaal. Selleks et sellise filtri leidmine oleks edukas, peab soovitud signaal olema sisendsignaali statistiliselt seotud. See tähendab, et soovitud signaal peab olema sisendsignaali abil ennustatav. Edasi leiab adaptiivne filter veasignaali, lahutades soovitud signaalist tema sisendsignaali abil formeeritud ennustuse. Vea signaali kasutatakse omakorda filtri koefitsientide korrigeerimiseks selliselt, et veasignaal muutuks teatud mõttes väiksemaks.

Selles õppematerjalis kasutame läbivalt järgnevaid tähistusi:

x – sisendsignaal,

y – adaptiivse filtri väljundsignaal,

d – soovitud signaal,

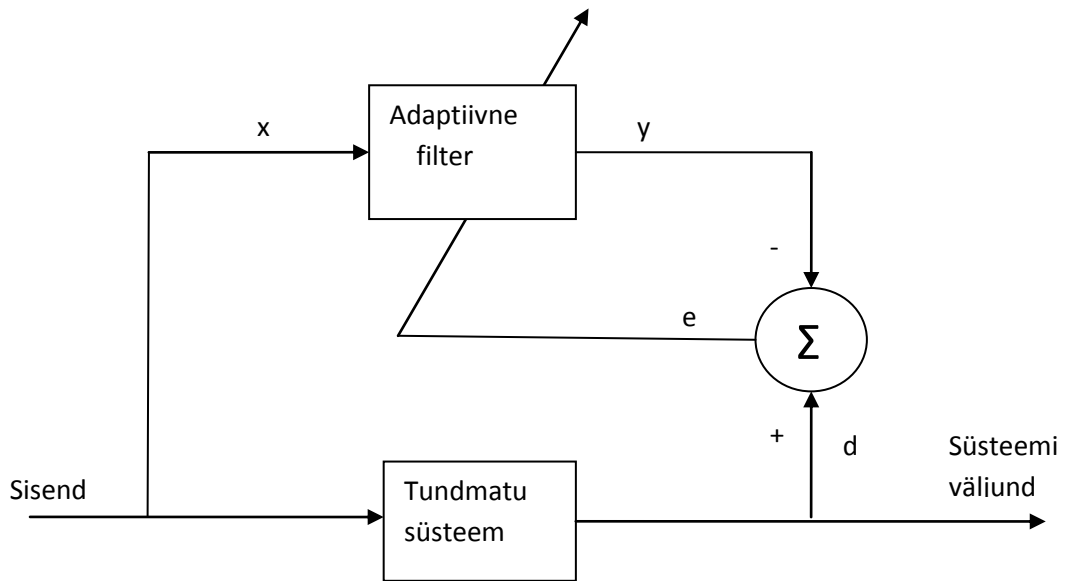
$e = d - y$, – veasignaal.

Võib eristada nelja põhilist adaptiivsete filtrite rakenduse tüüpi

1. Süsteemi identifitseerimine. Vaata Joonis 1. Selles rakenduses on meil võimalik jälgida teatud süsteemi sisendit ja väljundit ning adaptiivse filtri ülesandeks on leida jälgitava süsteemi lineaarne mudel st sellise lineaarse filtri koefitsiendid, mis rakendatuna sisendile annavad soovitud signaalile võimalikult sarnase filtri väljundsignaali. Sellisel juhul on süsteemi väljundsignaaliks tundmatu süsteemi väljundsignaal ja adaptiivse filtri töö tulemuseks on mudeli koefitsiendid.

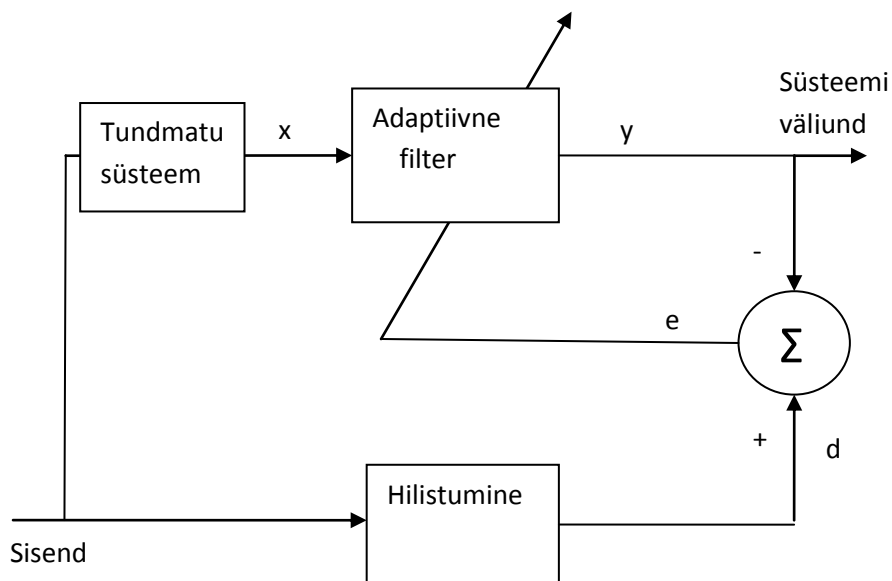
Antud konfiguratsiooni kasutatakse ka sidesüsteemides vajalikes kajakõrvaldajates. Näiteks on akustilised kajakad vaja kõrvaldada juhul, kui telefonikõne tegemiseks kasutatakse eemal asuvat valjuhääldit ning mikrofoni. Sellisel juhul on tundmatuks süsteemiks ruumi impulsskarakteristik ning adaptiivne filter formeerib signaali, mis sarnaneb ruumi kjadega

mikrofoni sisendis. Süsteemi väljundiks on veasignaal, kus kajad on maha surutud.



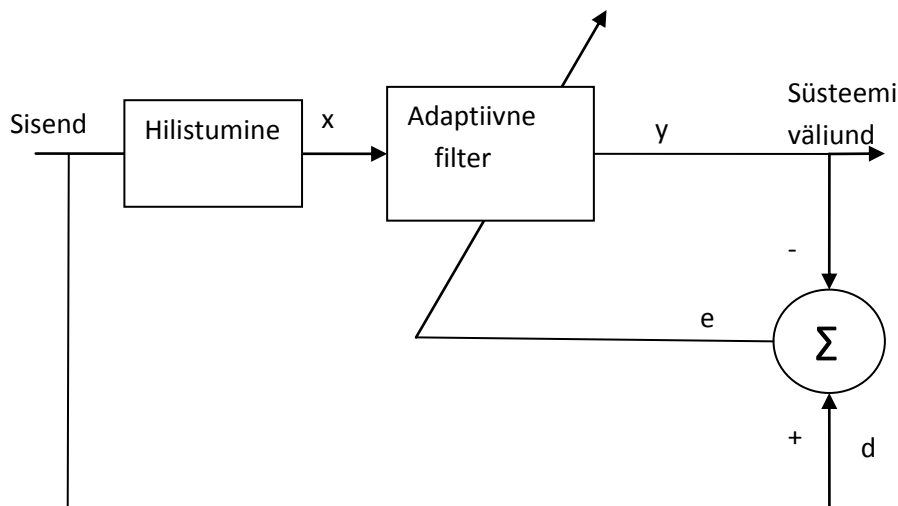
Joonis 1. Adaptiivse filtri rakendus süsteemi identifitseerimiseks

2. Pöördmudeli leidmine. Ka selles rakenduses on meil võimalik jälgida teatud süsteemi sisendit ja väljundit, kuid erinevalt eelmisest rakendusest on siin adaptiivse filtri ülesandeks leida jälgitava süsteemi lineaarne pöördmudel. Kui tundmatu süsteem on lineaarne, siis oleks tema ideaalse pöördmudeli sageduskarakteristik võrdne üks jagatud tundmatu süsteemi sageduskarakteristikuga nii, et süsteemi ja pöördmudeli järjestikühenduse sageduskarakteristik oleks ühikimpulss. Enamikul juhtudel on selles rakenduses vaja adaptiivse filtri kausaalsuse tagamiseks kasutada soovitud signaalina hilistatud sisendsignaali. Selline konfiguratsioon on näiteks laialdaselt kasutusel sidekanali ekvalaiserites.



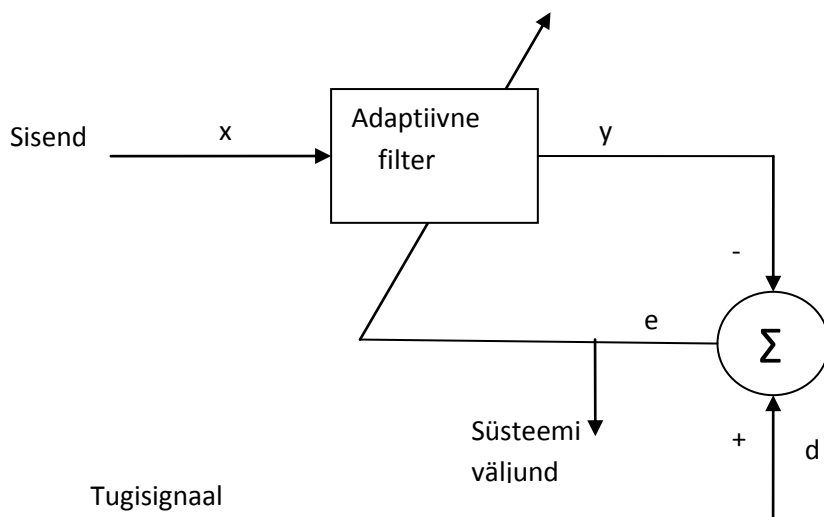
Joonis 2. Adaptiivse filtri kasutamine pöördmudeli leidmiseks

3. Lineaarne ennustamine. Selles rakenduses on adaptiivne filter kasutusel juhusliku signaali ennustajana. Soovitud signaalina kasutatakse signaali väärtust antud ajahetkel ning sisendsignaalina tema väärtusi eelnevatel ajahetkedel. Adaptiivne filter toimib selles rakenduses kui filter, mis ennustab signaali tulevasi väärtusi minevikuväärtustest nii, et väljundis ilmub signaali lineaarne ennustus. Selline ennustamine on seda edukam, mida korreleeritumad on signaali väärtused erinevatel ajahetkedel. Sõltuvalt rakendusest võib süsteemi väljundina olla kasutusel kas adaptiivse filtri väljund või veasignaali. Skeem on kasutusel ennustavates kõne koodrites nagu ADPCM (Adaptive Differential Pulse Code Modulation) [6], aga ka näiteks signaalide spektraalanalüüsis [8].



Joonis 3. Adaptiivse filtri kasutamine signaali lineaarse ennustajana

4. Häire allasurumine. Selles rakenduses kasutatakse adaptiivset filtrit tundmatu häire allasurumiseks teatud huvipakkuvas signaalis. Häireks võib antud kontekstis olla kas segav(ad) signaal(id) allikast, mis meile huvi ei paku, või müra või nende segu. Rakendus eeldab mitme sensori olemasolu süsteemis. Adaptiivse filtri soovitud signaalina kasutatakse signaali sensorilt, mis registreerib huvipakkuva signaali ja häire segu. Adaptiivse filtri sisendsignaalina kasutatakse signaali sensoritelt, mis on paigutatud selliselt, et nad ei registreeri meile huvipakkuvat signaali, vaid ainult häireid. Sellisel juhul formeerib adaptiivne filter signaali, mis sarnaneb häirekomponendiga soovitud signaalis ning lahutab selle huvipakkuva signaali ja häirete segust, vähendades häirete osakaalu veasignaalis. Skeem on kasutusel mürade allasurumisel näiteks aktiivsetes müra summutavates kõrvklappides, aga ka 50 Hz sagedusega toitevõrgu häirete vähendamisel EKG signaalide salvestamisel. Samuti on skeem kasutusel paljudes antennivõresid kasutavates rakendustes.



Joonis 4. Adaptiivse filtri kasutamine häirete allasurumiseks

3. Wieneri filtrid

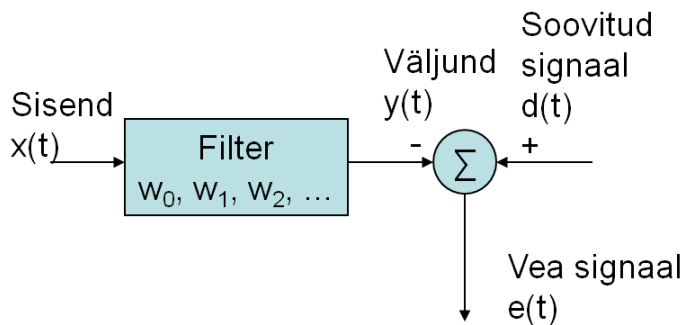
Selles peatükis leiavad käsitlemist optimaalsed diskreetses ajas opereerivad filtrid, mida kutsutakse Wieneri filtriteks.

Wieneri filter ei ole iseenesest adaptiivne, kuid ta kujutab endast sellegipoolest väga olulist vahendit adaptiivsete filtrite mõistmisel, seades piirid sellele, millise täpsuse on adaptiivne filter suuteline saavutama väljakujunenud režiimis st pärast seda, kui adaptiivne algoritm on koondunud. Nimelt jäävad tüüpilise adaptiivse filtri koefitsiendid pärast esialgset koondumisperioodi suhteliselt stabiilseks, hakates liikuma väikese amplituudiga teatud tasakaalupunkti läheduses. Selleks tasakaalupunktiks ongi paljude adaptiivsete algoritmide jaoks Wieneri filter. Tüüpiliselt koosneb adaptatsiooniviga Wieneri filtri veast ning adaptatsioonist tingitud lisaveast.

Selles peatükis huvitab meid konstantsete koefitsientidega (kaaludega) lineaarne filter, mille kaalud on valitud andmaks filtri väljundsignaalile soovitud omadusi. Soovitud omaduseks on veasignaali viimine nii väikseks kui võimalik teatud statistilises mõttes. Praktikas osutub mugavaimaks minimeerida veasignaali ruutkeskmist väärtust, kuna sel juhul on kriteeriumi sõltuvus filtri koefitsientidest ruutfunktsioon, millel on üks miinimum ja seda miinimumi on meeldivalt lihtne leida. Oma käsitluses vaatleme ainult Wieneri filtreid, mis opereerivad diskreetses ajas, seega huvitavad meid diskreetsed Wieneri filtrid. Peale selle piirdume me filtritega, mille väljundsignaal avaldub sisendsignaali väärtuste kaalutud summana

$$y(n) = \sum_{k=0}^{\infty} w_k^* x(n-k), \quad n = 0, 1, 2, \dots \quad (3.1)$$

kus w_k on k -s filtrikoefitsient ning tärn ülemise indeksina tähistab kompleksarvu kaaskompleksi. Valem (3.1) sisaldab lõpmatut summat ning sobib seega nii lõpmatu kui ka lõpliku impulsskarakteristikuga filtrite kirjeldamiseks. Esimesed tulemused saamegi sellise üldise lõpmatu impulsskarakteristikuga filtri jaoks, kuid suuremas osas käesolevast õppevahendist piirdume me lõpliku impulsskarakteristikuga filtrite vaatlemisega. Seda põhjusel, et lõpliku impulsskarakteristikuga filter on garanteeritult stabiilne, samal ajal kui lõpmatu impulsskarakteristikuga filtrid kasutavad tagasisidet ning nende stabiilsuse tagamiseks peab rakendama spetsiaalseid võtteid. Tulenevalt sellest on lõpliku impulsskarakteristikuga adaptiivsed filtrid suhteliselt lihtsamad kui lõpmatu impulsskarakteristikuga adaptiivsed filtrid.



Joonis 5. Wieneri filter

Kokkuvõtlikult on meie vaadeldav Wieneri filter diskreetse ajas toimiv lineaarne filter, mis formeerib diskreetsest sisendsignaalist $x(n)$ diskreetse väljundsignaali $y(n)$ selliselt, et filtri väljundsignaal oleks ruutkeskmise vea mõttes võimalikult sarnane soovitud signaalile $d(n)$. Teiste sõnadega, Wieneri filter on filter, mille kaalud w_i on valitud selliselt, et minimeerida vea signaali $e(n)$ ruutkeskmist väärtust $E[e^2(n)]$. Joonis 5 kujutab Wieneri filtrit skemaatiliselt.

Ortogonaalsuse printsiip

Me alustame Wieneri filtri tuletamist ortogonaalsuse printsiibi käsitlemisest. Olgu meil signaal diskreetidega $x(0), x(1), x(2), \dots$ ning filter koefitsientidega w_0, w_1, w_2, \dots (vaata Joonis 5). Filtri väljundsignaal avaldub konvolutsioonisummana

$$y(n) = \sum_{k=0}^{\infty} w_k^* x(n-k), \quad n = 0, 1, 2, \dots \quad (3.2)$$

Me eeldame, et filtri sisendsignaal ning soovitud signaal on vastastikku laias mõttes statsionaarsed nullise keskvaartusega juhuslikud protsessid. Me soovime leida selliseid filtrikoefitsiente, et filtri väljundsignaal $y(n)$ oleks vaadeldav soovitud signaali $d(n)$ hinnanguna ning vaatleme selle hinnangu viga

$$e(n) = d(n) - y(n) \quad (3.3)$$

Selleks et leida parimad võimalikud filtrikoefitsiendid, minimeerime veasignaali ruutkeskmist väärtust

$$J = E[e(n)e^*(n)] = E[|e(n)|^2] \quad (3.4)$$

kus operaator $E[\cdot]$ tähistab matemaatilist ootust. Meie filtri sisendsignaali on kompleksne ja seega on üldjuhul ka filtrikoefitsiendid kompleksarvud. Tähistame filtrikoefitsientide reaali- ja imaginaariosad vastavalt a ning b -ga

$$w_k = a_k + jb_k, \quad k = 0, 1, 2, \dots \quad (3.5)$$

Edasi defineerime gradiendioperaatori, mille k -s element on esimest järku osatuletis filtri koefitsiendi reaali- ja imaginaariosade järgi

$$\nabla_k = \frac{\partial}{\partial a_k} + j \frac{\partial}{\partial b_k}, \quad k = 0, 1, 2, \dots \quad (3.6)$$

Seega on veasignaali ruutkeskmise väärtuse vastav gradient kompleksne vektor, mille k -s element avaldub kujul

$$\begin{aligned} \nabla_k J &= \frac{\partial}{\partial a_k} J + j \frac{\partial}{\partial b_k} J \\ &= E \left[\frac{\partial e(n)}{\partial a_k} e^*(n) + \frac{\partial e^*(n)}{\partial a_k} e(n) + \frac{\partial e(n)}{\partial b_k} j e^*(n) + \frac{\partial e^*(n)}{\partial b_k} j e(n) \right], \quad k = 0, 1, 2, \dots \end{aligned} \quad (3.7)$$

Neli eelmises valemis figureerivat osatuletist on leitavad, kasutades valemeid (3.3) ja (3.2) ning võttes vajalikud tuletised kui

$$\begin{aligned} \frac{\partial e(n)}{\partial a_k} &= -x(n-k) \\ \frac{\partial e^*(n)}{\partial a_k} &= -x^*(n-k) \\ \frac{\partial e(n)}{\partial b_k} &= -jx(n-k) \\ \frac{\partial e^*(n)}{\partial b_k} &= -jx^*(n-k) \end{aligned} \quad (3.8)$$

Asendades leitud tuletised ruutkeskmise vea gradiendi avaldisse (3.7), saame tulemuseks

$$\nabla_k J = -2E[x(n-k)e^*(n)] \quad k = 0, 1, 2, \dots \quad (3.9)$$

Ruutkeskmise vea, mis on reaalne funktsioon, miinimumi leidmiseks võrdsustame tema gradiendi nulliga, mis annab

$$E[x(n-k)e^*(n)] = 0, \quad k = 0, 1, 2, \dots \quad (3.10)$$

Seega on vajalik ja piisav tingimus J miinimumiks kõigi sisendi x diskreetide, mis on kasutuses hinnangu leidmiseks, ja hindamisvea $e(n)$ ortogonaalsus. Seda tulemust kutsutaksegi ortogonaalsuse printsiibiks.

Järeldus. Vaatleme optimaalse filtri väljundsignaali ning veasignaali vahelist korrelatsiooni. Kasutades filtri väljundsignaali arvutamiseks valemit (3.2), saame

$$E[y(n)e^*(n)] = E\left[\sum_{k=0}^{\infty} w_k^* x(n-k)e^*(n)\right] = \sum_{k=0}^{\infty} w_k^* E[x(n-k)e^*(n)] \quad (3.11)$$

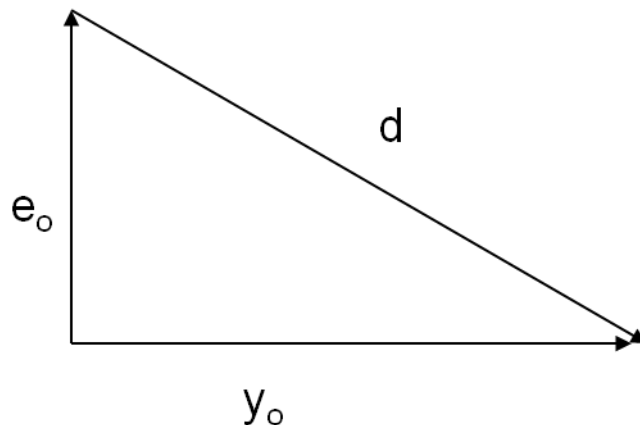
Seega, tähistades vähimruutude mõttes optimaalse filtri poolt produtseeritud signaale alumise indeksiga o (väljundsignaali $y_o(n)$ ja veasignaali $e_o(n)$) ning, kasutades ortogonaalsuse printsiipi (3.10), saame tulemuseks

$$E[y_o(n)e_o^*(n)] = 0 \quad (3.12)$$

Leitud ortogonaalsuse printsiibi järelduse võime seega sõnastada:

Vähima ruutkeskmise vea mõttes optimaalse filtri väljundsignaal on ortogonaalne tema poolt produtseeritud veasignaaliga.

Seda tulemust võib kujutada geomeetriliselt, nagu näitab Joonis 6. Filtri väljundsignaal y_o on soovitud signaali d parim hinnang minimaalse ruutkeskmise vea mõttes. Samas on filtri väljundsignaal, mis on lineaarne kombinatsioon sisendsignaali x diskreetidest, ortogonaalne veasignaaliga. Me võime seega öelda, et filtri väljundsignaal on soovitud signaali projektsioon sisendsignaali poolt määratud alamruumi. See tähendab vektorruumi, mis koosneb kõigist vektoritest, mida on võimalik moodustada sisendvektorite lineaarse kombinatsioonina.



Joonis 6. Optimaalse filtri väljundsignaali, soovitud signaali ja veasignaali vaheline suhe

Vähima veasignaali ruutkeskmise väärtuse, mis saavutatakse optimaalsete filtrikoefitsientide korral, saame nüüd, kasutades signaalide vahelist seost (3.3) ja äsjaleitud ortogonaalsuse printsiibi järeldust (3.12), avaldada kujul

$$J_{\min} = E[e_o(n)^2] = E[d(n)^2] - E[y_o(n)^2] \quad (3.13)$$

Wiener-Hopfi võrrandid

Järgnevalt leiame avaldise optimaalsete filtrikoefitsientide jaoks. Ortogonaalsuse printsiibi (3.10) põhjal võime kirjutada

$$E\left[x(n-k)\left(d^*(n)-\sum_{i=0}^{\infty}w_{oi}x^*(n-i)\right)\right]=0, \quad k=0,1,2,\dots \quad (3.14)$$

kus w_{oi} on optimaalse filtri i -s koefitsient. Avades sulud, saame

$$\sum_{i=0}^{\infty}w_{oi}E[x(n-k)x^*(n-i)]=E[x(n-k)d^*(n)], \quad k=0,1,2,\dots \quad (3.15)$$

Vaatleme ülaltoodud avaldises sisalduvaid matemaatilisi ootusi. Neist esimese argumendiks on sisendsignaali diskreetide ajahetkedel $n-k$ ja $n-i$ kaaskompleksne korrutis. Matemaatiline ootus sellest korrutisest kujutab endast sisendsignaali autokorrelatsiooni kohal $i-k$. Seega tähistame

$$E[x(n-k)x^*(n-i)]=r(i-k) \quad (3.16)$$

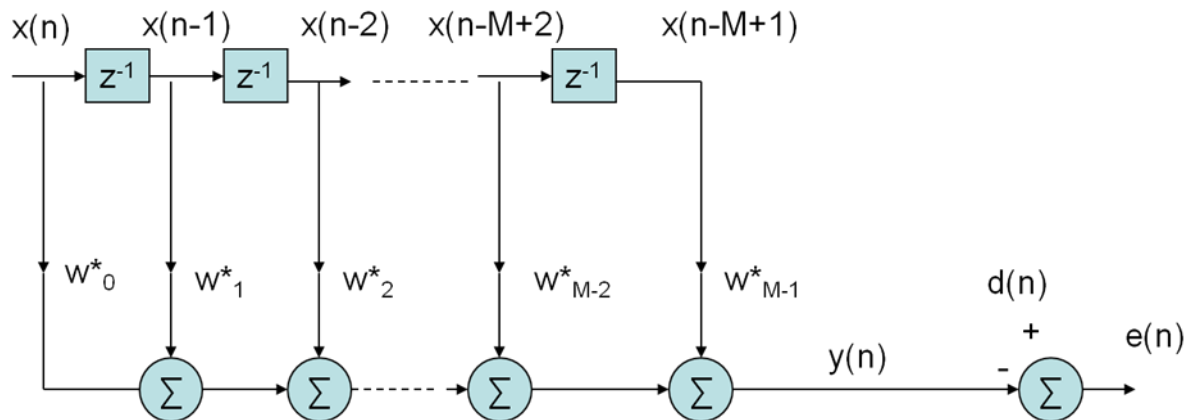
Teine matemaatiline ootus kujutab endast sisendsignaali kohal $n-k$ ja soovitud signaali kohal n ristkorrelatsiooni. Seega on ta ristkorrelatsioon kohal $-k$.

$$E[x(n-k)d^*(n)]=p(-k) \quad (3.17)$$

Asendades ülaltoodud tähistused valemisse (3.15), saame lõpmata paljudest võrranditest koosneva võrrandisüsteemi

$$\sum_{i=0}^{\infty}w_{oi}r(i-k)=p(-k), \quad k=0,1,2,\dots \quad (3.18)$$

Saadud võrrandisüsteem seob vähima ruutkeskmise vea mõttes optimaalse filtri koefitsiente sisendsignaali autokorrelatsiooni ning sisendsignaali ja soovitud signaali vahelise ristkorrelatsiooniga. Võrrandisüsteem on tuntud Wiener-Hopfi võrrandite nime all. Äsjaleitud kujul on Wiener-Hopfi võrrandisüsteemi lahendamine keerukas, kuna ta sisaldab lõpmatult palju tundmatuid w_{oi} . Ülesanne lihtsustub tunduvalt, kui piirdume hinnangu leidmisel lõpliku impulsskarakteristikuga filtriga. Olgu filtri koefitsientide arv M , nagu näitab Joonis 7. Plokid, mis on tähistatud sümbooliga z^{-1} , väljendavad signaali hilistumist ühe diskreedi võrra.



Joonis 7. Lõpliku impulsskarakteristikuga filter

Wiener-Hopfi võrrandisüsteem sellise filtri jaoks on filtri lõplikust ulatusest tulenevalt lõplik

$$\sum_{i=0}^{M-1} w_{oi} r(i-k) = p(-k), \quad k = 0, 1, 2, \dots, M-1 \quad (3.19)$$

Saamaks filtreerimisoperatsiooni kirjeldamiseks kompaktsemat tähistust, võime me filtri mälus ajahetkel n talletatud M sisendsignaali diskreeti koondada veeruvektorisse

$$\mathbf{x}(n) = [x(n) \quad x(n-1) \quad \dots \quad x(n-M+1)]^T \quad (3.20)$$

kus ülemine indeks T tähistab transponeerimist. Samuti kogume veeruvektorisse otsitavad filtrikoefitsiendid

$$\mathbf{w} = [w_0 \quad w_1 \quad \dots \quad w_{M-1}]^T \quad (3.21)$$

Filtri väljundsignaali ajahetkel n saame nüüd kompaktselt avaldada kui vektorite skalaarkorrutise

$$y(n) = \sum_{i=0}^{M-1} w_i^* x(n-i) = \mathbf{w}^H \mathbf{x}(n) \quad (3.22)$$

kus ülemine indeks H tähistab Hermiti transponeerimist

$$[a_0 + jb_0 \quad a_1 + jb_1 \quad \dots \quad a_{M-1} + jb_{M-1}]^H = \begin{bmatrix} a_0 - jb_0 \\ a_1 - jb_1 \\ \vdots \\ a_{M-1} - jb_{M-1} \end{bmatrix}. \quad (3.23)$$

Edasi koondame maatriksisse \mathbf{R} Wiener-Hopfi võrrandisüsteemis (3.19) figureerivad autokorrelatsiooni väärtused

$$\mathbf{R} = E[\mathbf{x}(n)\mathbf{x}^H(n)] = \begin{bmatrix} r(0) & r(1) & \dots & r(M-1) \\ r^*(1) & r(0) & \dots & r(M-2) \\ \vdots & \vdots & \ddots & \vdots \\ r^*(M-1) & r^*(M-2) & \dots & r(0) \end{bmatrix} \quad (3.24)$$

Maatriksit \mathbf{R} kutsutakse sisendsignaali autokorrelatsiooni maatriksiks. Autokorrelatsioonimaatriks on konstruktsiooni põhjal Toeplitzi maatriks. See tähendab, et tema mis tahes diagonaalil paiknevad elemendid on kõik omavahel võrdsed (peadiagonaalil on kõik elemendid võrdsed $r(0)$ -ga, esimesel diagonaalil $r(1)$ -ga jne).

Peale selle koondame vektorisse \mathbf{p} Wiener-Hopfi võrrandisüsteemis (3.19) figureerivad sisendsignaali ja soovitud signaali ristkorrelatsiooni väärtused

$$\mathbf{p} = E[\mathbf{x}(n)d^*(n)] \quad (3.25)$$

Nende tähistustega saame Wiener-Hopfi võrrandisüsteemi kirjutada maatrikskujul kui

$$\mathbf{R}\mathbf{w}_o = \mathbf{p} \quad (3.26)$$

Ning eeldusel, et maatriks \mathbf{R} ei ole singulaarne, avalduvad optimaalse filtri koefitsiendid kui

$$\mathbf{w}_o = \mathbf{R}^{-1}\mathbf{p} \quad (3.27)$$

kus ülemine indeks -1 tähistab pöördmaatriksi leidmise operatsiooni. Seega on meil vaja optimaalse lõpliku impulsskarakteristikuga filtri leidmiseks teada sisendsignaali $\mathbf{x}(n)$ autokorrelatsioonimaatriksit \mathbf{R} ning sisendsignaali $\mathbf{x}(n)$ ja soovitud signaali $d(n)$ vahelist ristkorrelatsiooni vektorit \mathbf{p} .

Vea pind

Leiame nüüd võrrandi, mis seob veasignaali ruutkeskmist väärtust filtri koefitsientidega. Seda selleks, et paremini mõista, kuidas sõltub veasignaali ruutkeskmise väärtus filtri koefitsientide kaugusest nende optimaalsetest väärtustest. Selline seos osutub vägagi kasulikuks, kui me edaspidi uurime adaptiivsete filtrite omadusi. Veasignaali võime, kasutades avaldisi (3.3) ja (3.22), kirjutada kujul

$$e(n) = d(n) - \mathbf{w}^H \mathbf{x}(n) \quad (3.28)$$

Kasutades ülaltoodut ning tähistusi (3.24) ja (3.25), saame veasignaali ruutkeskmise väärtuse kirja panna kui

$$\begin{aligned} J &= E[e(n)e^*(n)] = E[(d(n) - \mathbf{w}^H \mathbf{x}(n))(d^*(n) - \mathbf{x}^H(n)\mathbf{w})] \\ &= E[d(n)d^*(n) - d(n)\mathbf{x}^H(n)\mathbf{w} - \mathbf{w}^H \mathbf{x}(n)d^*(n) + \mathbf{w}^H \mathbf{x}(n)\mathbf{x}^H(n)\mathbf{w}] \\ &= \delta_d^2 - \mathbf{p}^H \mathbf{w} - \mathbf{w}^H \mathbf{p} + \mathbf{w}^H \mathbf{R}\mathbf{w} \end{aligned} \quad (3.29)$$

Signaali autokorrelatsiooni maatriks on garanteeritult mittenegatiivselt määratud, aga praktikas peaaegu alati positiivselt määratud maatriks, kuna praktikas esineb enamasti signaalile lisanduv mürakomponent. Seega on maatriks \mathbf{R} praktikas peaaegu alati pööratav ja seetõttu võime valem (3.29) esitada ekvivalentselt kui

$$J(\mathbf{w}) = \delta_d^2 - \mathbf{p}^H \mathbf{R}^{-1} \mathbf{p} + (\mathbf{w} - \mathbf{R}^{-1} \mathbf{p})^H \mathbf{R} (\mathbf{w} - \mathbf{R}^{-1} \mathbf{p}) \quad (3.30)$$

Ülaltoodud avaldise kehtivust on lihtne kontrollida, kui avame sulud ja koondame sarnased liikmed.

Kuna optimaalse filtri koefitsiendid on antud valemiga (3.27), muutub eeltoodud avaldise viimane liige optimaalse filtri korral nulliks ning seega avaldub minimaalne võimalik veasignaali ruutkeskmine väärtus kui

$$\min_{\mathbf{w}} J(\mathbf{w}) = J_{\min} = \delta_d^2 - \mathbf{p}^H \mathbf{R}^{-1} \mathbf{p} \quad (3.31)$$

Kasutades sellist tähistust, võime võrrandi (3.30) ümber kirjutada kui

$$J(\mathbf{w}) = J_{\min} + (\mathbf{w} - \mathbf{R}^{-1} \mathbf{p})^H \mathbf{R} (\mathbf{w} - \mathbf{R}^{-1} \mathbf{p}) = J_{\min} + (\mathbf{w} - \mathbf{w}_o)^H \mathbf{R} (\mathbf{w} - \mathbf{w}_o) \quad (3.32)$$

Seega koosneb veasignaali ruutkeskmine väärtus oma minimaalse võimaliku väärtuse ning ühe ruutvormi, mis on seotud filtrikoefitsientide kaugusega nende optimaalsest väärtusest summast. Eelnevalt öeldu paremaks mõistmiseks kasutame korrelatsioonimaatriksi lahutust omaväärtuste ning omavektorite järgi [4, 15]. Kuna autokorrelatsioonimaatriksil on Hermiti sümmeetria, võime ta lahutada osadeks kujul

$$\mathbf{R} = \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^H \quad (3.33)$$

kus $\mathbf{\Lambda}$ on diagonaalne omaväärtuste maatriks $\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_M)$ ning maatriksi \mathbf{Q} veerud on korrelatsioonimaatriksi \mathbf{R} omavektorid. Kuna kompleksne maatriks \mathbf{Q} koosneb maatriksi \mathbf{R} omavektoritest, mis on ortogonaalsed, on maatriks \mathbf{Q} unitaarne maatriks. Seega

$$\mathbf{Q}^H \mathbf{Q} = \mathbf{I} \quad (3.34)$$

ning

$$\mathbf{Q}^{-1} = \mathbf{Q}^H \quad (3.35)$$

Asendades lahutuse (3.33) valemisse (3.32), saame

$$J(\mathbf{w}) = J_{\min} + (\mathbf{w} - \mathbf{w}_o)^H \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^H (\mathbf{w} - \mathbf{w}_o) \quad (3.36)$$

Edasi tähistame lineaarselt (ortogonaalse maatriksiga \mathbf{Q}) teisendatud koefitsientide vektori ning optimaalse koefitsientide vektori vahe vektori kui

$$\mathbf{v} = \mathbf{Q}^H (\mathbf{w} - \mathbf{w}_o) \quad (3.37)$$

Selle tähistusega avaldub (3.36) kui

$$J(\mathbf{w}) = J_{\min} + \mathbf{v}^H \mathbf{\Lambda} \mathbf{v} \quad (3.38)$$

Kuna Λ on diagonaalne maatriks, ei sisalda ülaltoodud avaldise teine liidetav risttegeureid, ja seega lahutub ülaltoodud ruutvorm lihtsaks summaks

$$J(\mathbf{w}) = J_{\min} + \mathbf{v}^H \Lambda \mathbf{v} = J_{\min} + \sum_{k=1}^M v_k^* \lambda_k v_k = J_{\min} + \sum_{k=1}^M \lambda_k |v_k|^2 \quad (3.39)$$

Saadud avaldist kutsutakse vea pinna kanooniliseks kujuks. Tasub tähele panna, et vektorid \mathbf{v} määravad vea pinna teljed.

Näide

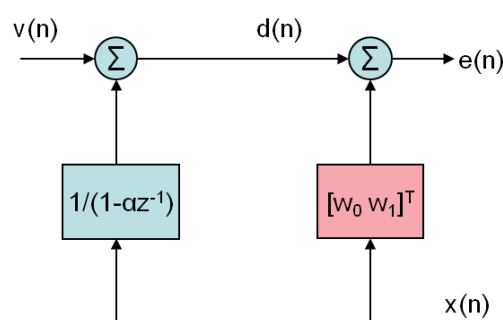
Joonis 8 kujutab kahe koefitsiendiga filtrit $[w_0 \ w_1]^T$, mis on kasutusel tundmatu süsteemi identifitseerimiseks. On vaja leida minimaalse ruutkeskmise vea mõttes optimaalsed koefitsiendid (Wieneri filter) ja vähima ruutkeskmise vea võimsus, kui $x(n)$ ja $v(n)$ on sõltumatud valged mürad ja $|\alpha| < 1$.

Wieneri filtri konstrueerimiseks on meil vaja sisendsignaali korrelatsioonimaatriksit ning sisendsignaali ja soovitud signaali vahelise ristkorrelatsiooni vektorit. Kuna sisendsignaali x on valge müra, on tema korrelatsioonimaatriks ühikmaatriks, mille peadiagonaali elemendid on võrdsed müra võimsusega

$$\mathbf{R} = \begin{bmatrix} \delta_x^2 & 0 \\ 0 & \delta_x^2 \end{bmatrix} \quad (3.40)$$

Soovitud signaal koosneb filtreeritud sisendsignaali ja valge müra $v[n]$ summast. Seega võime väljundsignaali z teisenduse kirjutada kui

$$D(z) = \frac{X(z)}{1 - \alpha z^{-1}} + V(z) \quad (3.41)$$



Joonis 8. Süsteemi identifitseerimine kahe koefitsiendiga filtri abil

Ajavallas väljendab (3.41) rekursiivset seost soovitud signaali diskreetide vahel

$$\begin{aligned}
d(n) &= x(n) + \alpha d(n-1) + v(n) \\
d(n-1) &= x(n-1) + \alpha d(n-2) + v(n) \\
d(n) &= x(n) + \alpha x(n-1) + \alpha^2 d(n-2) + v(n)
\end{aligned} \tag{3.42}$$

...

seada rekursiooni võime jätkata lõputult ning seega avaldub soovitud signaal sisendsignaali kaudu järgmise lõpmatu summana:

$$d(n) = \sum_{i=0}^{\infty} \alpha^i x(n-i) + v(n) \tag{3.43}$$

Arvestades x -i ning v statistilist sõltumatust ning pannes tähele, et lõpmatu summa eelnevas valemis on geomeetriline progressioon, saame soovitud signaali võimsuseks

$$\delta_d^2 = E[d^2(n)] = \sum_{i=0}^{\infty} \alpha^{2i} \delta_x^2 + \delta_v^2 = \frac{\delta_x^2}{1-\alpha^2} + \delta_v^2 \tag{3.44}$$

Edasi leiame sisendsignaali ning soovitud signaali vastastikuse korrelatsioonivektori, kasutades valemit (3.43) ning protsesside x ja v statistilist sõltumatust

$$\mathbf{p} = E \left[d(n) \begin{bmatrix} x(n) \\ x(n-1) \end{bmatrix} \right] = \begin{bmatrix} \delta_x^2 \\ \alpha \delta_x^2 \end{bmatrix} \tag{3.45}$$

Otsitav Wieneri filter on nüüd leitav valemist (3.27) kui

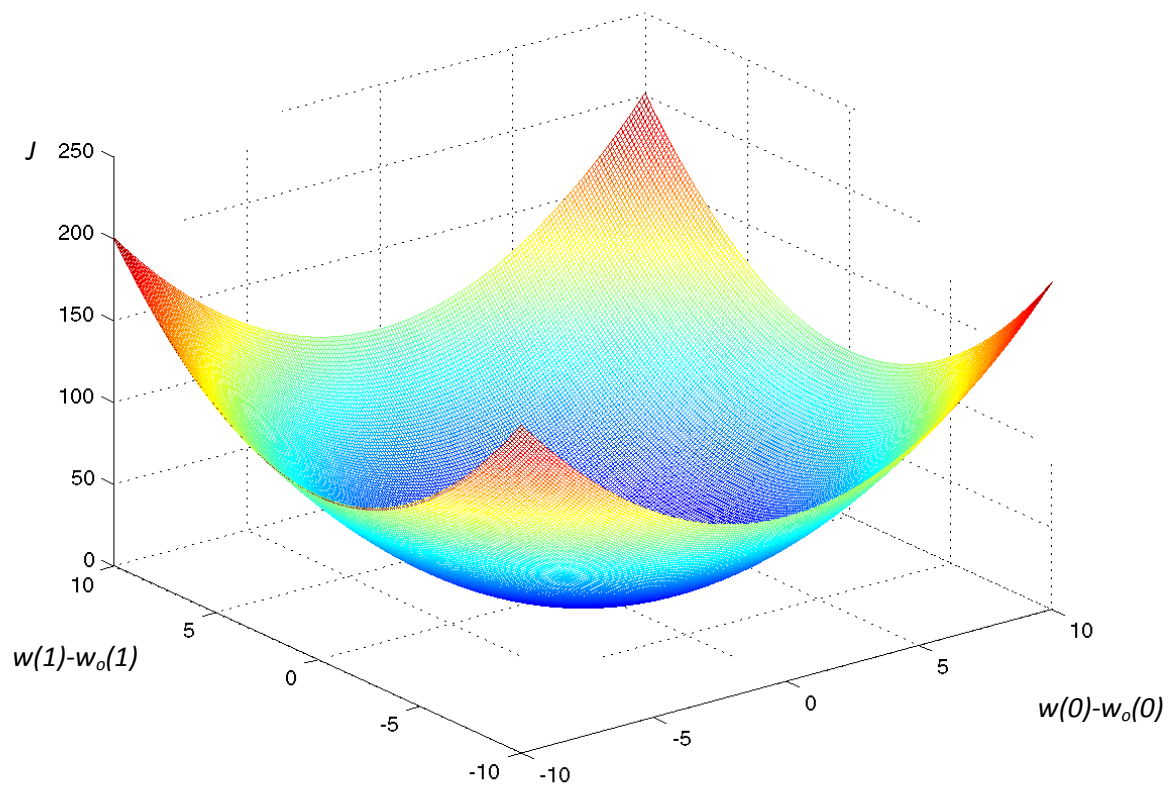
$$\mathbf{w}_o = \mathbf{R}^{-1} \mathbf{p} = \begin{bmatrix} \delta_x^2 & 0 \\ 0 & \delta_x^2 \end{bmatrix}^{-1} \begin{bmatrix} \delta_x^2 \\ \alpha \delta_x^2 \end{bmatrix} = \begin{bmatrix} 1 \\ \alpha \end{bmatrix} \tag{3.46}$$

Tasub tähele panna, et Wieneri filtri koefitsientideks on $x(n)$ ning $x(n-1)$ kordajad valemist (3.42). Kõik suurema viitega sisendsignaali komponendid soovitud signaalis on jäänud modelleerimata. See on ka loomulik, arvestades, et sisendsignaali on valge müra ja seega erinevate sisendsignaali diskreetide vaheline korrelatsioon on võrdne nulliga.

Vähima ruutkeskmise vea võimsuse saame, kasutades valemit (3.31)

$$\begin{aligned}
J_{\min} &= \delta_d^2 - \mathbf{p}^H \mathbf{R}^{-1} \mathbf{p} = \frac{\delta_x^2}{1-\alpha^2} + \delta_v^2 - \delta_x^2 \begin{bmatrix} 1 & \alpha \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ \alpha \end{bmatrix} = \\
&= \frac{\delta_x^2}{1-\alpha^2} + \delta_v^2 - \delta_x^2 (1 + \alpha^2)
\end{aligned} \tag{3.47}$$

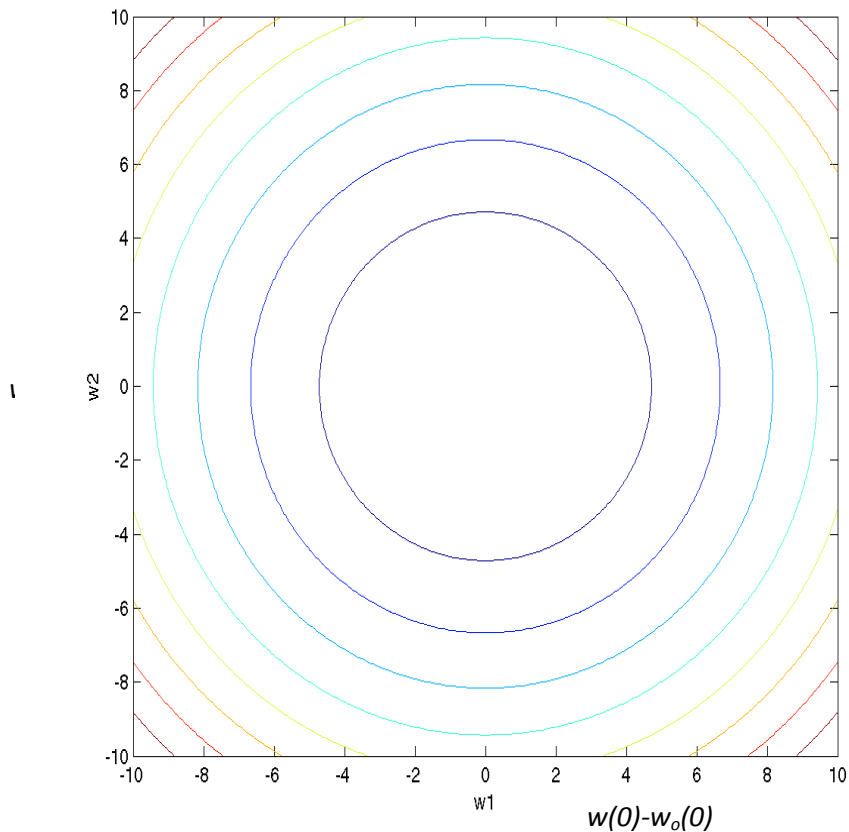
Leitud avaldise kaks esimest liidetavat väljendavad soovitud signaali võimsust avaldisest (3.44) ning kolmas liidetav vähendab seda avaldises (3.44) sisalduva lõpmatu summa kahe esimese liikme võrra. Ruutkeskmise viga sisaldab seega kõigi nende komponentide võimsust, mida kahe koefitsiendiga filter ei ole võimeline modelleerima. Selleks on nii müra v võimsus, kui ka kõigist sisendsignaali komponentidest, välja arvatud $x(n)$ ning $x(n - 1)$, tulenev võimsus.



Joonis 9. Vea pind

Joonis 9 kujutab näites toodud ülesandele vastavat vea pinda (3.32) funktsioonina filtri koefitsientide kõrvalekaldest nende optimaalsetest väärtustest, mis on antud valemiga (3.46). Loomulikult on vea pinnaks sellises koordinaadistikus paraboloid, mille miinimum paikneb punktis $[0, 0]$ ehk selliste filtri koefitsientide väärtuste juures, kus viga on võrdne nulliga.

Mõningates situatsioonides, nagu me edaspidi veendume, on mugavam vaadata vea pinda otse ülevalt. Joonis 10 toobki ära taolise vea pinna kujutise, kasutades selleks samakõrgusjooni, mis, arvestades vea pinna paraboloidset kuju, on üldjuhul ellipsid. Antud ülesandes, kus sisendsignaaliks on valge müra, mille korrelatsioonimaatriks on diagonaalne, on samakõrgusjooned küll ringid.



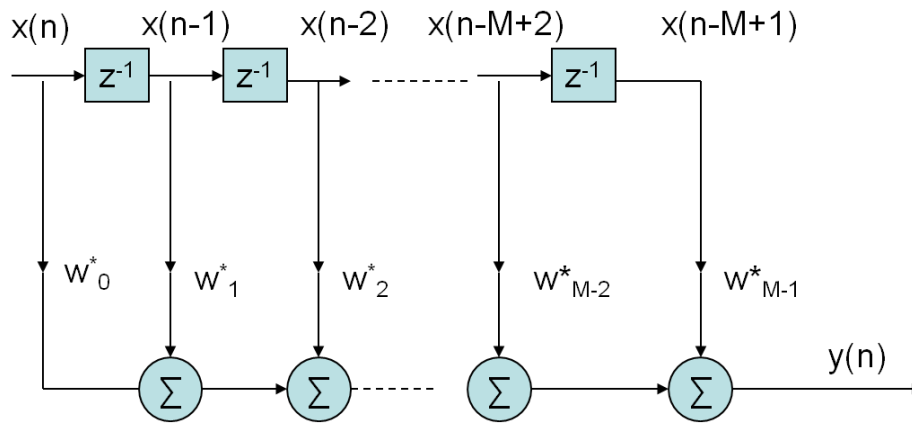
Joonis 10. Vea pind esitatuna samakõrgusjoone abil

Lineaarse piiranguga ülesanne

Siani oleme vaadelnud ülesandeid, kus on olnud vaja minimeerida veasignaali ruutkeskmist väärtust lahendile mingeid piiranguid seadmata. Kusjuures veasignaali on defineeritud soovitud signaali ja tegeliku filtri väljundsignaali vahena. Teatud juhtudel on aga soovitatav või vajalik konstrueerida filtreid, mis peale mõne signaali ruutkeskmise väärtuse minimeerimise rahuldavad ka teatud lisapiiranguid. Näiteks võime vajada filtrit, mis minimeerib väljundsignaali keskmist võimsust, tagades samal ajal mingi kindla võimenduse mingil kindlal sagedusel. Selles peatükis vaatlemegi, kuidas taolisi ülesandeid lahendada.

Joonis 11 kujutab lõpliku impulsskarakteristikuga filtrit, mille väljundsignaal avaldub kujul

$$y(n) = \sum_{k=0}^{M-1} w_k^* x(n-k), \quad (3.48)$$



Joonis 11. Lõpliku impulsskarakteristikuga filter

Erijuhul, kui filtri sisendsignaali on kompleksne eksponent

$$x(n) = e^{j\omega n}, \quad (3.49)$$

avaldub filtri väljundsignaal kujul

$$y(n) = e^{j\omega n} \sum_{k=0}^{M-1} w_k^* e^{-j\omega k} \quad (3.50)$$

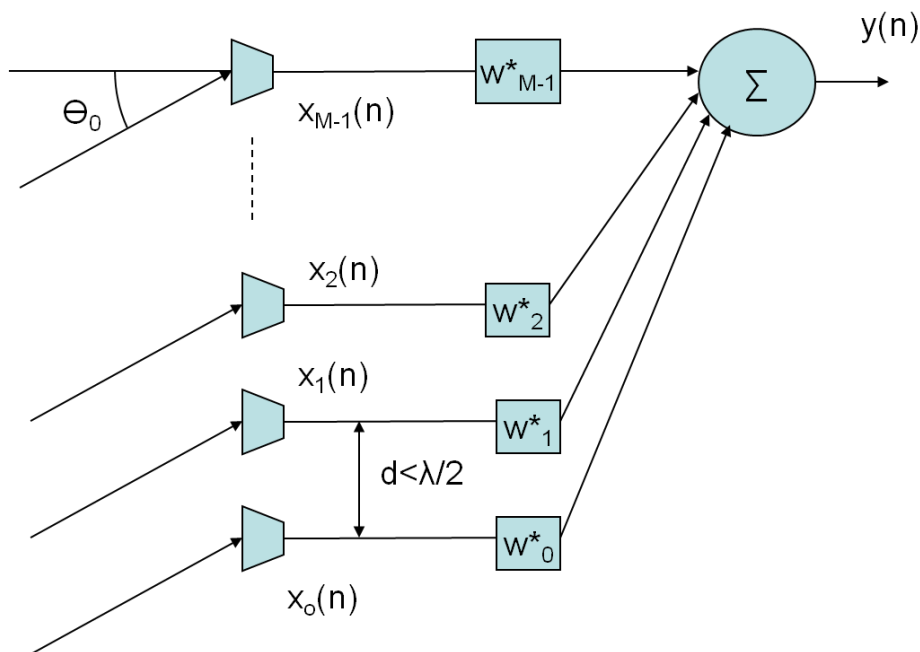
kus ω on diskreetimissageduse suhtes normaliseeritud ringsagedus. Paneme tähele, et eeltoodud valemis esinev summa on filtri sageduskarakteristik (diskreetne Fourier' teisendus impulsskarakteristikust).

Piiranguga optimeerimisülesande, mida me soovime lahendada, võime nüüd formuleerida kui: leida optimaalsed filtri koefitsiendid w_k , mis minimeerivad väljundi ruutkeskmist väärtust piiranguga

$$\sum_{k=0}^{M-1} w_k^* e^{-j\omega_0 k} = g \quad (3.51)$$

Seega nõuame, et filtri võimendus sagedusel ω_0 oleks g sõltumata filtrile rakendatud signaalidest. Eelnevas valemis rahuldab ringsagedus võrratust $-\pi < \omega \leq \pi$ ning g on mingi kompleksarv.

Ülalöeldu iseloomustab piiranguga ajalise filtreerimise. Võimalik on formuleerida ka analoogiline ruumilise filtreerimise ülesanne.



Joonis 12. Lineaarne antennivõre

Joonis 12 kujutab M antennist koosnevat lineaarset antennivõret, kus antennid on paigutatud võrdsete vahedega d , mis on väiksem kui pool lainepikkust λ . Piirang antennide vahekaugusele tagab võimaluse määrata langemisnurka üheselt ja on analoogiline Nyquisti diskreetimisteoreemiga aja vallas. Kõikide antennide väljundsignaalid korrutatakse komplekssete koefitsientidega ning summeeritakse. Langeagu sellisele antennivõrele kitsaribaline tasapinnaline laine nurga θ_0 alt. Kui tasapinnaline laine ei lange antennivõrele otse, vaid teatud nurga alt, jõuab signaal võret moodustavatele antennidele erinevatel ajahetkedel, põhjustades faasinihkeid erinevate antennide abil registreeritud signaalide vahel.

Elektriline nurk, mis väljendab signaalide faasinihet teineteise suhtes, on seotud ruumilise nurgaga kui

$$\phi = \frac{2\pi d}{\lambda} \sin(\theta) \tag{3.52}$$

Antennivõre väljundsignaal on suunast θ_0 tuleva laine puhul

$$y(n) = x_0(n) \sum_{k=0}^{M-1} w_k^* e^{-jk\phi_0}, \tag{3.53}$$

ning meie ülesanne võib olla formuleeritud järgmiselt: leida optimaalsed filtri koefitsiendid w , mis minimeerivad väljundi y ruutkeskmist väärtust piiranguga, et teatud suunast θ_0 langev laine jõuaks väljundisse võimendusega g ,

$$\sum_{k=0}^{M-1} w_k^* e^{-jk\phi_0} = g \tag{3.54}$$

Elektriline nurk rahuldab võrratust $-\pi < \phi \leq \pi$ ning g on kompleksarv.

Võrreldes ajalise ülesande formuleeringut ruumilise ülesande formuleeringuga näeme, et nad langevad kokku, kuigi kirjeldavad täiesti erinevaid probleeme. Mõlemal korral tuleb lahendada sama piiranguga optimeerimisülesanne.

$$\min J = E \left[\sum_{k=0}^{M-1} \sum_{i=0}^{M-1} w_k^* x^*(k) w_i x(i) \right] = \sum_{k=0}^{M-1} \sum_{i=0}^{M-1} w_k^* w_i r(i-k), \quad \text{piiranguga} \quad \sum_{k=0}^{M-1} w_k^* e^{-j\phi_0 k} = g \quad (3.55)$$

Lahendame selle ülesande Lagrange'i määramata kordajate meetodil. Kõigepealt defineerime reaalse kriteeriumifunktsiooni

$$J = \underbrace{\sum_{k=0}^{M-1} \sum_{i=0}^{M-1} w_k^* w_i r(i-k)}_{\text{väljundvõimsus}} + \text{Re} \left[\underbrace{\lambda^* \left(\sum_{k=0}^{M-1} w_k^* e^{-j\phi_0 k} - g \right)}_{\text{lineaamepiirang}} \right] = \mathbf{w}^H \mathbf{R} \mathbf{w} + \text{Re} \left[\lambda^* (\mathbf{w}^H \mathbf{s}(\phi_0) - g) \right] \quad (3.56)$$

kus vektorit \mathbf{s} , mis sisaldab ainsana informatsiooni piirangus esineva sageduse või suuna kohta, kutsutakse tüürvektoriks

$$\mathbf{s}(\phi_0) = \begin{bmatrix} 1 & e^{-j\phi_0} & \dots & e^{-j(M-1)\phi_0} \end{bmatrix}^T \quad (3.57)$$

Edasi leiame kriteeriumifunktsiooni gradiendi. Selleks et lugeja harjuks maatrikstähistustega, mida me ülejäänud õppevahendis valdavalt kasutame, esitame mõned järgnevad valemid kirja panduna nii summade abil (vasakul) kui ka maatrikskujul (paremal). On kerge veenduda, et vasak- ning parempoolne kirja pilt on ekvivalentsed.

$$\nabla_k J = 2 \sum_{i=0}^{M-1} w_i r(i-k) + \lambda^* e^{-j\phi_0 k} \quad \nabla J = 2 \mathbf{R} \mathbf{w} + \lambda^* \mathbf{s}(\phi_0) \quad (3.58)$$

Võrdsustades gradiendi nulliga, saame (taas kahel kujul)

$$\sum_{i=0}^{M-1} w_{oi} r(i-k) = -\frac{\lambda^*}{2} e^{-j\phi_0 k}, \quad k = 0, 1, \dots, M-1; \quad \mathbf{R} \mathbf{w}_o = -\frac{\lambda^*}{2} \mathbf{s}(\phi_0) \quad (3.59)$$

seega saime (eeldusel, et korrelatsioonimaatriks ei ole singulaarne) lahendiks

$$\mathbf{w}_o = -\frac{\lambda^*}{2} \mathbf{R}^{-1} \mathbf{s}(\phi_0) \quad (3.60)$$

Saadud lahendis figureerib tundmatu Lagrange'i kordaja λ , mille leiame, kasutades piirangut (3.54), mis maatrikskujul avaldub kui

$$\mathbf{w}_o^H \mathbf{s}(\phi_0) = g \quad (3.61)$$

Võttes ülesande lahendist (3.60) Hermiti transponaadi ning korrutades seejärel võrrandi mõlemad pooled paremalt tüürvektoriga \mathbf{s} , saame

$$\mathbf{w}_o^H \mathbf{s}(\phi_0) = -\frac{\lambda}{2} \mathbf{s}^H(\phi_0) \mathbf{R}^{-H} \mathbf{s}(\phi_0) \quad (3.62)$$

Nüüd paneme tähele, et võrrandite (3.61) ja (3.62) vasakud pooled on võrdsed, seega peavad olema võrdsed ka paremad pooled, ja me võime määramata kordaja λ avaldada võrrandite parematest pooltest kujul

$$\lambda = -\frac{2g}{\mathbf{s}^H(\phi_0)\mathbf{R}^{-1}\mathbf{s}(\phi_0)} \quad (3.63)$$

Asendades leitud λ eelnevalt filtrikoefitsientide jaoks saadud lahendisse (3.60), on tulemuseks

$$\mathbf{w}_o = -\frac{g^*}{\mathbf{s}^H(\phi_0)\mathbf{R}^{-1}\mathbf{s}(\phi_0)}\mathbf{R}^{-1}\mathbf{s}(\phi_0) \quad (3.64)$$

Saadud lahendi esimene – murruline liige – on kompleksne konstant, millele järgneb sisendsignaali autokorrelatsioonimaatriksi pöördmaatriksi ning tüürvektori korrutis. Seega on saadud lahend vägagi sarnane Wieneri filtriga (3.27), vaid selle vahega, et Wieneri filtris figureeriv sisendsignaali ning soovitud signaali ristkorrelatsioonivektor on asendunud teatud konstandiga korrutatud tüürvektoriga. Üks selle sarnasuse võimalikke interpretatsioone on, et piiranguga ülesandes etendab tüürvektor analoogilist rolli soovitud signaaliga piiramata ülesandes. Mõlemad väljendavad meie arusaama sellest, mida „hea“ filter tegema peaks.

Kasutades eelnevas piirangut $g = 1$, saame antennivõre, mis valitud suunast saabuvat signaali ei muuda, küll aga minimeerib kõigist teistest suundadest saabuvaid signaale. Sellist antennivõre töötlust kutsutakse vähima dispersiooni ning moonutusteta karakteristikuga suunadiagrammi formeerijaks (Minimum Variance Distortionless Response Beamformer).

$$g = 1 \Rightarrow \mathbf{w}_o = -\frac{\mathbf{R}^{-1}\mathbf{s}(\phi_0)}{\mathbf{s}^H(\phi_0)\mathbf{R}^{-1}\mathbf{s}(\phi_0)} \quad (3.65)$$

Sellise suunadiagrammi formeerija minimaalne väljundvõimsus avaldub ruutvormina

$$J_{\min} = \mathbf{w}_o^H \mathbf{R} \mathbf{w}_o \quad (3.66)$$

ning, asendades eelnevasse optimaalsed koefitsiendid avaldisest (3.65), saame

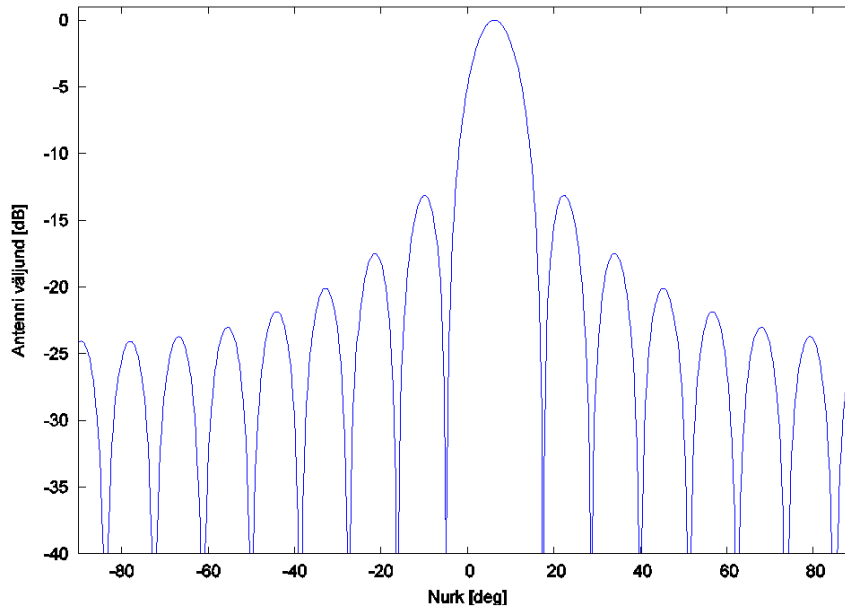
$$J_{\min} = \frac{\mathbf{s}^H(\phi_0)\mathbf{R}^{-1}\mathbf{R}\mathbf{R}^{-1}\mathbf{s}(\phi_0)}{\mathbf{s}^H(\phi_0)\mathbf{R}^{-1}\mathbf{s}(\phi_0)\mathbf{s}^H(\phi_0)\mathbf{R}^{-1}\mathbf{s}(\phi_0)} = \frac{1}{\mathbf{s}^H(\phi_0)\mathbf{R}^{-1}\mathbf{s}(\phi_0)} \quad (3.67)$$

Selline optimaalne kiireformeeriija laseb elektrilise nurgaga ϕ_0 sisendkomponente läbi muutmata kujul, surudes samal ajal maha müra ning signaale, mis saabuvad muudest suundadest. Seega on J_{\min} oma olemuselt nurga ϕ_0 suunalt saabuva signaali dispersioon (võimsus).

Huvitav on märkida, et vaadeldes eelnevat avaldist elektrilise nurga ϕ funktsioonina saame ühe tuntud võimsusspektri hindamise algoritmidest, nimelt Caponi võimsusspektri hinnangu [12].

$$S_{MVDR} = \frac{1}{\mathbf{s}^H(\phi)\mathbf{R}^{-1}\mathbf{s}(\phi)} \quad (3.68)$$

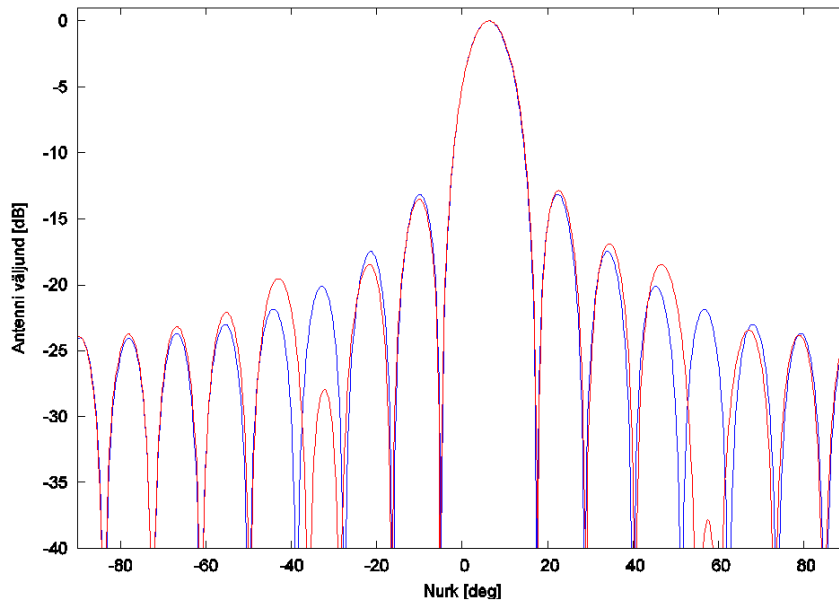
Näide. Vaatleme vähima dispersiooni ja muutmata karakteristikuga kiireformeerijat, mis kasutab 16 antenni, mille vahel on $d/\lambda=0.2$. Asugu soovitud allikas nurga all $\varphi_0=10^\circ$ ning olgu ainsaks häireks vastuvõtjate valge omamüra võimsusega 0.001. Erinevate kanalite omamürad on statistiliselt sõltumatud ning võrdsete võimsustega.



Joonis 13. Antennivõre suunadiagramm

Joonis 13 kujutab antud tingimustel leitud antennivõre suunadiagrammi. Nagu jooniselt nähtub, on vähima dispersiooni ja muutmata karakteristikuga kiireformeerija pealeht suunatud soovitud allikale, mis asub nurga $\varphi_0=10^\circ$ all, surudes samal ajal maha signaale kõigist ülejäänud suundadest.

Lisades häired, mis saavad suundadest $+60^\circ$ ja -30° ning jättes kõik muud tingimused samaks, saame järgmise suunadiagrammi.



Joonis 14. Antennivõre suunadiagrammid koos ning ilma häireteta

Sinise joonega antud suunadiagramm pärineb eelmiselt jooniselt ning näitab, kuidas vähima dispersiooni ja muutmata karakteristikuga kiireformeerija suunadiagramm formeerub ilma häireteta situatsioonis. Punane joon näitab, kuidas formeerub suunadiagramm häirete olemasolul antennivõrele langevas signaalis. Suunadiagrammi pealeht on endiselt suunatud soovitud signaalile, mis saabub nurga 10° alt. Peale selle on algoritm formeerinud sügavad miinimumid häirete saabumise suundadesse $+60^\circ$ ja -30° , minimeerides nii häirete mõju väljundsignaalis.

4. Kiireima languse meetod

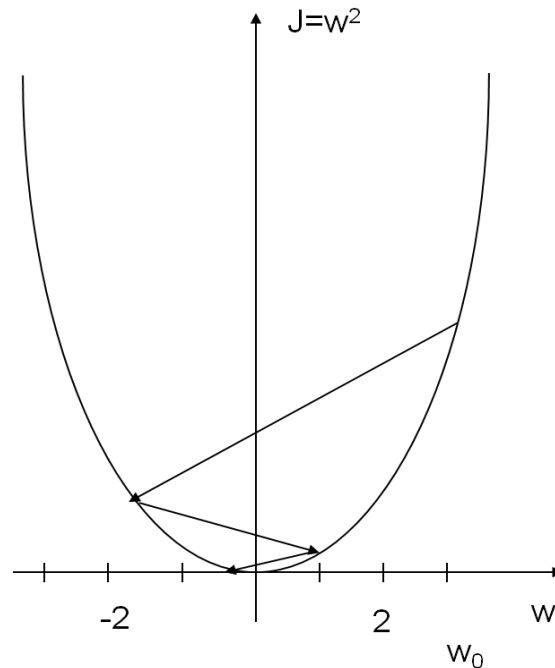
Selles peatükis käsitleme ühel eakal optimeerimistehnikal põhinevat algoritmi, mis on aluseks vea pinna gradiendil põhinevate adaptiivsete algoritmide töö mõistmisel. Meetod on oma olemuselt rekursiivne ning selle kasutamisel lähenetakst optimeerimisprobleemi lahendamisele järk-järguliste väikeste sammudega. Rakendatuna Wieneri filtri probleemile annab meetod meile võimaluse läheneda optimaalsele lahendamisele ilma sisendsignaali korrelatsioonimaatriksit pööramata.

Garanteeritud globaalsesse optimumi koondumiseks eeldab meetod, et meie kriteeriumifunktsioonil on üks miinimum. Wieneri filtri puhul me nägime, et vea pind on ruutfunktsioon tundmatustest filtrikoefitsientidest ning seega on antud eeldus täidetud. Kiireima languse meetodi idee on äärmiselt lihtne ning võime algoritmi formuleerida järgmiselt:

1. Alustada suvalise esialgse koefitsientide vektoriga $\mathbf{w}(0)$, (lisainformatsiooni puudusel võetakse tavaliselt algvektoriks nullvektor).
2. Leida gradientvektor st ruutkeskmise vea $J(n)$ tuletis $\mathbf{w}^*(n)$ järgi.
3. Leida uus koefitsientide vektor \mathbf{w} , astudes sammu gradiendi vastassuunas.

4. Tagasi sammule 2.

Joonis 15 kujutab endast kiireima languse meetodi illustratsiooni lihtsa ühe koefitsiendiga filtri puhul. Kuna ülesanne on ühemõõtmeline, on kriteeriumifunktsioon lihtsalt parabool, mis sõltub ainult ühest argumentist, milleks on otsitav filtrikoefitsient w .

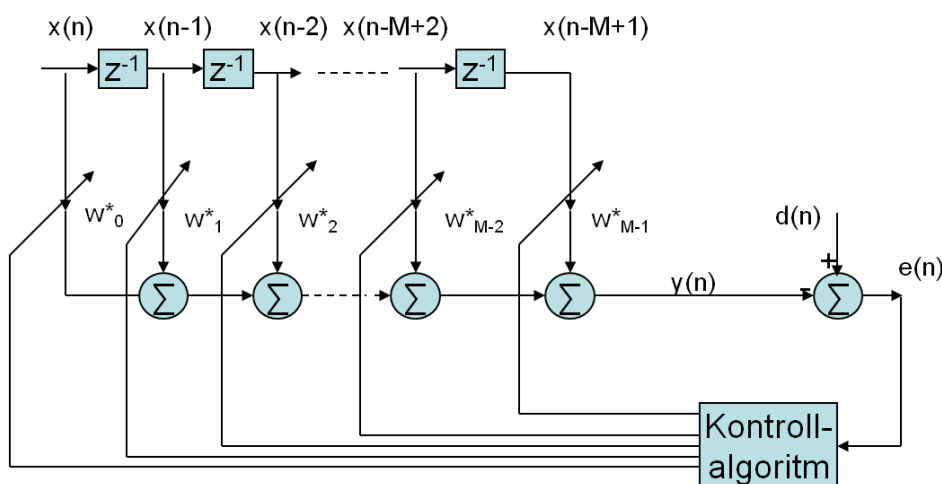


Joonis 15. Ühemõõtmelise kiireima languse meetodi illustratsioon

Kriteeriumifunktsiooni $J = w^2$ tuletis on $dJ/dw = 2w$. Olgu meie algühenduseks $w_0 = 3$. Kriteeriumi tuletis punktis w_0 on võrdne $dJ/dw |_{w_0} = 6$, mis on positiivne suurus. Seega järgmise lähendi leidmiseks peame astuma sammu negatiivses suunas. Eeltoodud lihtne algoritm ei anna mingeid juhiseid sammu pikkuse valikuks, aga oletame, et me astume sammu pikkusega 5, jõudes seeläbi punkti $w_1 = -2$. Selles punktis leiame jällegi kriteeriumi tuletise, milleks on $dJ/dw |_{w_1} = -4$. Kuna tuletis on negatiivne, astume järgmise lähendi leidmiseks sammu positiivses suunas, kusjuures sammu pikkuseks olgu nüüd 3, jõudes punkti $w_2 = 1$. Leiame taas kriteeriumi tuletise $dJ/dw |_{w_2} = 2$, mis on positiivne, seega tuleb meil jätkata sammuga negatiivses suunas. Seda protseduuri võime jätkata lõpmatult, leides üha paremaid lähendeid kriteeriumi miinimumpunkti koordinaadile. See eeldab muidugi, et meie astunud sammud on mõistlikult lühikesed, nii et me üle pinna põhja liiga kaugele ei hüppaks ning miinimumile lähenemise asemel temast hoopis kaugeneks.

Kiireima languse meetodi rakendus Wieneri filtrile

Joonis 16 kujutab lõpliku impulsskarakteristikuga adaptiivset filtrit sisendsignaaliga x , väljundsignaaliga y , soovitud signaaliga d ning koefitsientidega w .



Joonis 16. Lõpliku impulsskarakteristikuga adaptiivne filter

Vastavalt kiireima languse meetodile, tuleb uuendatud koefitsientide vektori leidmiseks igal iteratsioonil astuda samm gradiendile vastupidises suunas. Teiste sõnadega, uue koefitsientide vektori lähenduse saame liites vanale koefitsientide vektorile mingi konstandiga korrutatud ja miinusmärgiga võetud gradiendivektori. Selle matemaatiline üleskirjutus oleks

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{1}{2} \mu [-\nabla J(n)] \quad (4.1)$$

kus μ on sobivalt valitud konstant, mis peab kindlustama piisavalt lühikesed sammud tagamaks algoritmi koondumist.

Ruutkeskmine viga sammul n on vastavalt valemile (3.29)

$$J(n) = \delta_d^2 - \mathbf{p}^H \mathbf{w}(n) - \mathbf{w}^H(n) \mathbf{p} + \mathbf{w}^H(n) \mathbf{R} \mathbf{w}(n), \quad w_i = a_i + j b_i \quad (4.2)$$

kus \mathbf{p} on sisendsignaali ja soovitud signaali vaheline ristkorrelatsioonivektor ning \mathbf{R} on sisendsignaali autokorrelatsiooni maatriks. Filtri koefitsiendid on kompleksarvud $w_i = a_i + j b_i$. Kriteeriumifunktsiooni J gradiendi sammul n võime leida analoogiliselt eelmises peatükis toodud gradiendarvutustega (3.7), saades

$$\nabla J(n) = \begin{bmatrix} \frac{\partial J(n)}{\partial a_0(n)} + j \frac{\partial J(n)}{\partial b_0(n)} \\ \dots \\ \frac{\partial J(n)}{\partial a_{M-1}(n)} + j \frac{\partial J(n)}{\partial b_{M-1}(n)} \end{bmatrix} = -2\mathbf{p} + 2\mathbf{R}\mathbf{w}(n) \quad (4.3)$$

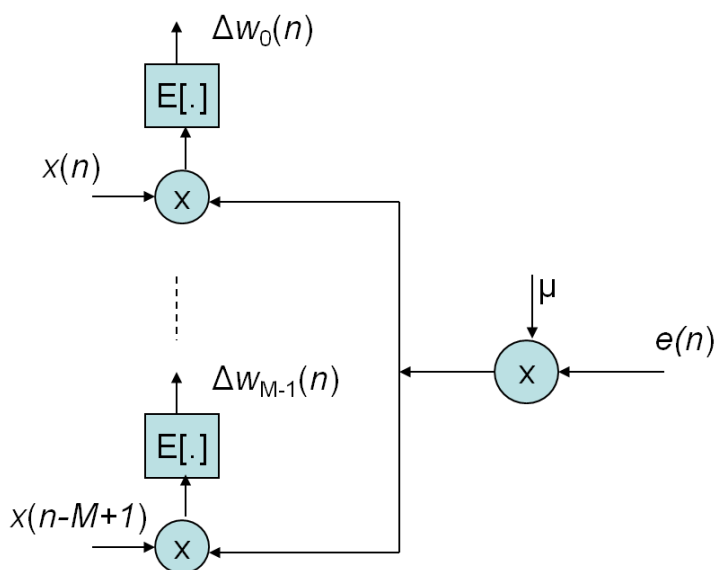
Asendades leitud gradiendi valemisse (4.1), saame kiireima languse meetodi rekursiooniks

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu [\mathbf{p} - \mathbf{R}\mathbf{w}(n)], \quad n = 0, 1, 2, \dots \quad (4.4)$$

Leitud valem väljendabki kiireima languse meetodit lõpliku impulsskajaga Wieneri filtri jaoks. Valemi kasutamine sellisel kujul eeldab, et sisendsignaali korrelatsioonimaatriks ning sisendsignaali ja soovitud signaali ristkorrelatsiooni vektor on teada. Kiireima languse meetodi rekursioonile saab anda ekvivalentse kuju, mis ei vaja eelnimetatud teavet

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu E[\mathbf{x}(n)e^*(n)], \quad n = 0, 1, 2, \dots \quad (4.5)$$

Valemite (4.4) ja (4.5) ekvivalentsuse tõestamine on jäetud lugejale iseseisva töö ülesandeks. Joonis 17 kujutab kiireima languse rekursiooni graafiliselt. Nagu nähtub jooniselt, võib kiireima languse meetodit ette kujutada kui M paralleelselt ühendatud ristkorrelatsiooni leidmise seadet, kus ristkorrelatsioonid leitaks veasignaali $e(n)$ hetkväärtuse vahel ühelt poolt ning M järjestikuse sisendsignaali x väärtuse vahel teiselt poolt. Kaalude juurdekasvud iteratsioonil n on joonisel tähistatud kui $\Delta w_i(n)$

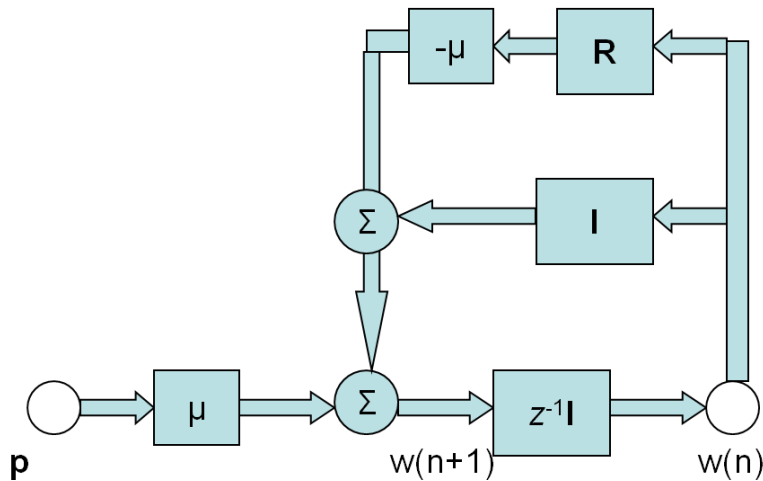


Joonis 17. Kiireima languse meetodi rekursioon

Kiireima languse meetodi stabiilsus

Eelnevalt mainisime, et kiireima languse meetodi abil funktsiooni miinimumi leidmiseks peab sammu pikkus μ olema sobivalt valitud. Käesolevas peatükis leiamegi tingimused, mida μ peab rahuldama selleks, et algoritm koolduks (oleks stabiilne).

Joonis 18 kujutab kiireima languse meetodit signaaligraafina. Nagu jooniselt näha, on kiireima languse meetod tagasisidestatud süsteem, mille stabiilsus sõltub suletud ahela võimendusest. \mathbf{I} tähistab ühikmaatriksit, mis võimendust ei mõjuta, ja seega sõltub stabiilsus ainult sammu suuruselt μ ning sisendsignaali korrelatsioonimaatriksist \mathbf{R} .



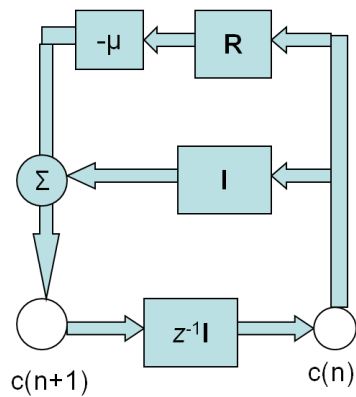
Joonis 18. Kiireima languse meetodi signaalgraaf

Defineerime koefitsientide vea vektori kui

$$\mathbf{c}(n) = \mathbf{w}(n) - \mathbf{w}_o \quad (4.6)$$

kus \mathbf{w}_o on kaalude vektor, mis vastab antud ülesande optimaalsele Wieneri lahendile. Kirjutame kiireima languse meetodi rekursiooni (4.4) ümber, kasutades koefitsientide vea vektorit

$$\begin{aligned} \mathbf{c}(n+1) &= \mathbf{w}(n) - \mathbf{w}_o + \mu(\mathbf{p} - \mathbf{R}\mathbf{w}(n)) \\ &= \mathbf{w}(n) - \mathbf{w}_o + \mu(\mathbf{R}\mathbf{w}_o - \mathbf{R}\mathbf{w}(n)) = (\mathbf{I} - \mu\mathbf{R})\mathbf{c}(n) \end{aligned} \quad (4.7)$$



Joonis 19. Kiireima languse meetodi signaalgraaf kasutades koefitsientide vea vektorit

Joonis 19 kujutab rekursiooni (4.7) signaalgraafina, rõhutades taas, et antud tagasisidestatud süsteemi stabiilsus sõltub ainult \mathbf{R} -st ja μ -st. Selleks, et leida stabiilsuse tingimused, lahutame maatriksi \mathbf{R} tema omavektorite ja omaväärtuste järgi kui

$$\mathbf{R} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H, \quad (4.8)$$

kus \mathbf{Q} on unitaarne omavektorite maatriks, mille veergudeks on maatriksi \mathbf{R} omavektorid. $\mathbf{\Lambda}$ on diagonaalne omaväärtuste maatriks, st diagonaalmaatriks, mille peadiagonaalil on kõik maatriksi \mathbf{R} omaväärtused samas järjekorras nagu omavektorid maatriksis \mathbf{Q} . Tulenevalt sellest, et \mathbf{R} on korrelatsioonimaatriks ja seega mittenegatiivselt määratud, on kõik tema omaväärtused reaalsed ning mittenegatiivsed. Asendades korrelatsioonimaatriksi lahutuse (4.8) valemisse (4.7), saame

$$\mathbf{c}(n+1) = (\mathbf{I} - \mu\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H)\mathbf{c}(n) \quad (4.9)$$

Kuna unitaarsete maatriksite omaduste põhjal on unitaarse maatriksi korrutis tema Hermiti transponaadiga võrdne ühikmaatriksiga, võime valemile (4.9) anda kuju

$$\mathbf{c}(n+1) = (\mathbf{Q}\mathbf{Q}^H - \mu\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H)\mathbf{c}(n) \quad (4.10)$$

ning seejärel tuua maatriksi \mathbf{Q} ning tema Hermiti transponaadi sulgudest välja, saades

$$\mathbf{c}(n+1) = \mathbf{Q}(\mathbf{I} - \mu\mathbf{\Lambda})\mathbf{Q}^H\mathbf{c}(n) \quad (4.11)$$

Järgmiseks korrutame võrrandi mõlemad pooli vasakult maatriksi \mathbf{Q} Hermiti transponaadiga, mis annab tulemuseks

$$\mathbf{Q}^H\mathbf{c}(n+1) = (\mathbf{I} - \mu\mathbf{\Lambda})\mathbf{Q}^H\mathbf{c}(n) \quad (4.12)$$

Nüüd viime sisse tähistuse

$$\mathbf{v} = \mathbf{Q}^H\mathbf{c}(n) = \mathbf{Q}^H(\mathbf{w} - \mathbf{w}_o) \quad (4.13)$$

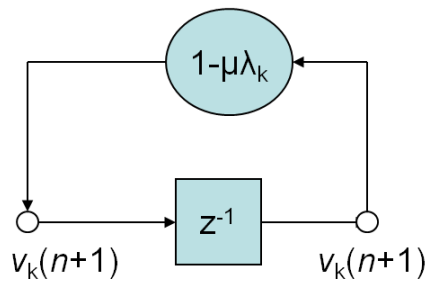
Vektor \mathbf{v} on lineaarselt (maatriksiga \mathbf{Q}) teisendatud koefitsientide vigade vektor. Kasutades seda tähistust, võime avaldise (4.12) kirjutada kujul

$$\mathbf{v}(n+1) = (\mathbf{I} - \mu\mathbf{\Lambda})\mathbf{v}(n) \quad (4.14)$$

Saadud avaldisest on näha, et kuna nii \mathbf{I} kui $\mathbf{\Lambda}$ on diagonaalmaatriksid, ei ole mis tahes vektori \mathbf{v} elemendi v_k areng sammult n sammule $n + 1$ seotud tema teiste elementidega ning võime valemit (4.14) lihtsustada, saades

$$v_k(n+1) = (1 - \mu\lambda_k)v_k(n) \quad (4.15)$$

Joonis 20 kujutab kiireima languse meetodi lihtsustunud signaaligraafi vastavalt valemile (4.15).



Joonis 20. Kiireima languse meetodi lihtsustatud signaaligraaf

Selleks et tagasisidestatud süsteem oleks stabiilne, peab tema võimendus olema väiksem kui üks. Rakendades seda teadmist äsja leitud struktuurile, on lihtne näha, et kiireima languse meetodi stabiilsuse tingimuseks on

$$-1 < 1 - \mu\lambda_k < 1, \quad \forall k \quad (4.16)$$

Kuna korrelatsioonimaatriksi omaväärtused on reaalsed ja mittenegatiivsed, saame eelnevast järgmise vajaliku ja piisava stabiilsuse tingimuse sammu suurusele

$$0 < \mu < \frac{2}{\lambda_{\max}} \quad (4.17)$$

kus λ_{\max} on sisendsignaali korrelatsioonimaatriksi \mathbf{R} suurim omaväärtus.

Valem (4.15) kujutab endast esimest järku homogeenset diferentsvõrrandit, mille lahendiks on

$$v_k(n) = (1 - \mu\lambda_k)v_k(n-1) = (1 - \mu\lambda_k)^n v_k(0) \quad (4.18)$$

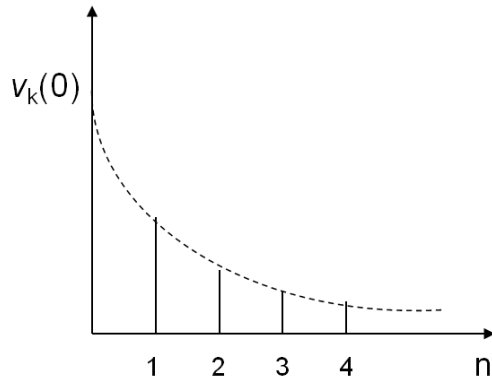
kus $v_k(0)$ on v_k algväärtus ajahetkel 0. Kuna kõik maatriksi \mathbf{R} omaväärtused on reaalsed ning mittenegatiivsed, ei esine v_k käitumises võnkumisi. Seega me näeme, et v_k areneb geomeetrilise reana sammuga $1 - \mu\lambda_k$. Oletame nüüd, et me oleme algoritmi sammu valinud vastavuses avaldisega (4.17). Siis on rea samm kõikide k -de korral väiksem kui 1 ning kõik moodid $v_k(n)$ lähenevad n kasvades monotoonselt nullile. Moodi all mõistame siin ühe omaväärtuse ja omavektori paari poolt määratud trajektoori. Selle geomeetrilise reaga saame sobitada eksponentsiaalse mähkija ajakonstandiga τ_k , valides

$$1 - \mu\lambda_k = \exp\left(-\frac{1}{\tau_k}\right) \quad (4.19)$$

Vastavalt eeltoodule on eksponentsiaalse mähkija ajakonstant arvutatav kui

$$\tau_k = \frac{-1}{\ln(1 - \mu\lambda_k)} \quad (4.20)$$

Ajakonstant τ_k on aeg, mille jooksul $v_k(n)$ väheneb e korda, kus e on naturaallogaritmide alus. Ülalöeldut illustreerib Joonis 21.



Joonis 21. $v_k(n)$ koondumine

Terve algoritmi koondumiskiiruse määrab kõikide moodide koostoime ja selle jaoks ei saa anda lihtsat valemit, nagu oli võimalik anda iga individuaalse moodi jaoks. Sellegipoolest saame väita, et kogu algoritmi koondumiskiirus jääb vähima ja suurima omaväärtuse poolt määratud moodide koondumiskiiruste vahele

$$\frac{-1}{\ln(1 - \mu\lambda_{\max})} \leq \tau_a \leq \frac{-1}{\ln(1 - \mu\lambda_{\min})} \quad (4.21)$$

kus τ_a on tervet algoritmi iseloomustav ajakonstant.

Võib tähele panna, et algoritmi koondumus sõltub sellest, kuidas erinevad teineteisest sisendsignaali korrelatsioonimaatriksi vähim ja suurim omaväärtus. Kui erinevus on väike, koondub algoritm suhteliselt kiiresti ja, vastupidi, kui erinevus on suur, koondub algoritm suhteliselt aeglaselt. Seda sellepärast, et algoritmi stabiilsuse tagamiseks vajaliku sammu pikkuse määrab suurim omaväärtus, ning kui ta on palju suurem kui ülejäänud väärtused, tingib see aeglase liikumise väikestele omaväärtustele vastavates suundades.

Edasi vaatleme algoritmi ruutkeskmise vea arengut. Kasutades vea pinna kanoonilist kuju (3.39), saame kirjutada

$$\begin{aligned} J(n) &= J_{\min} + (\mathbf{w} - \mathbf{w}_o)^H \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^H (\mathbf{w} - \mathbf{w}_o) = J_{\min} + \sum_{k=1}^M \lambda_k |v_k(n)|^2 \\ &= J_{\min} + \sum_{k=1}^M \lambda_k (1 - \mu\lambda_k)^{2n} |v_k(0)|^2 \end{aligned} \quad (4.22)$$

kus J_{\min} on Wieneri filtri poolt määratud vähim võimalik ruutkeskmise viga. Järelikult, kui kiireima languse meetod koondub (sobivalt valitud μ), siis läheneb avaldise (4.22) teine liidetav n kasvades nullile sõltumatult algtingimustest

$$\lim_{n \rightarrow \infty} J(n) = J_{\min} \quad (4.23)$$

Seega koondub kiireima languse meetod Wieneri filtri.

Algoritmi ruutkeskmist viga, mis on joonistatud aja funktsioonina, nimetatakse õppimiskõveraks. Kiireima languse meetodi õppimiskõver on algoritmi moode kirjeldavate eksponentide summa, kusjuures k -s omamood omab ajakonstanti

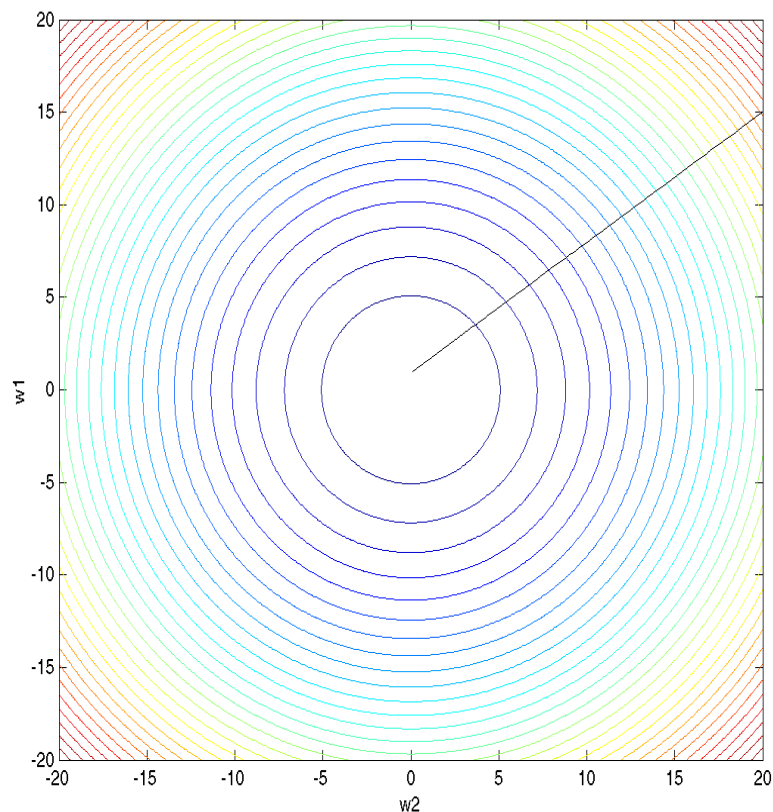
$$\tau_{k,mse} = \frac{-1}{2 \ln(1 - \mu \lambda_k)} \quad (4.24)$$

Kui sammu suurus μ on valitud väike, siis saame ülaltoodud ajakonstandi avaldise lähendada järgmiselt

$$\tau_{k,mse} \approx \frac{1}{2\mu\lambda_k}, \quad \mu \ll 1 \quad (4.25)$$

Näide

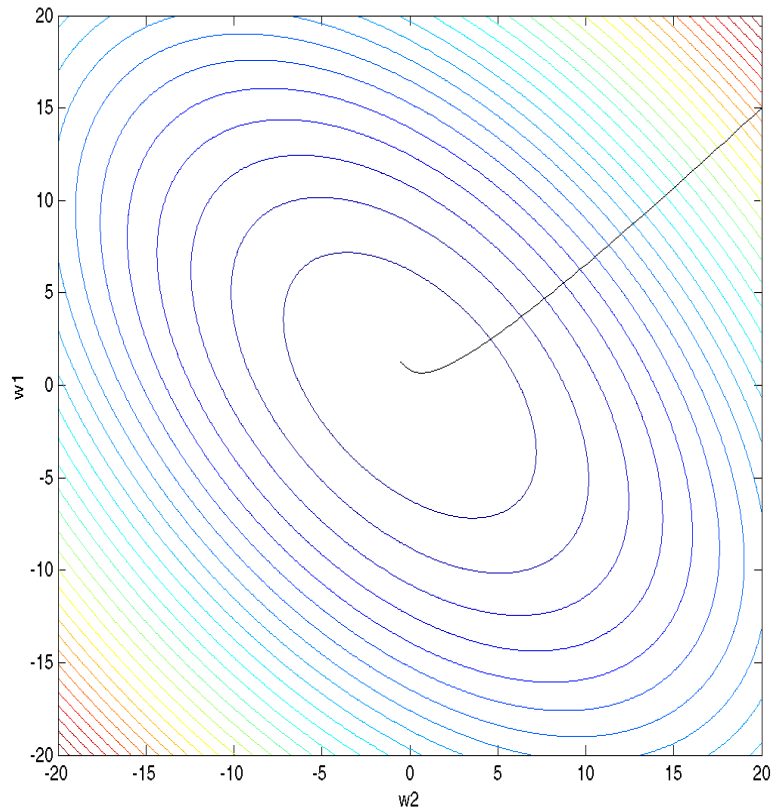
Selles näites vaatleme, kuidas näeb välja kiireima languse meetodi koondumise tee vea pinnal, kasutades näitlikkuse huvides lihtsat, kahe kaaluga filtrit. See võimaldab meil kujutada vea pinda koordinaadistikus, kus ühel teljel on ühe ja teisel teise filtri kaalukoefitsiendi väärtus.



Joonis 22. Kiireima languse meetodi kaalude vektori trajektoor valge sisendsignaali korral

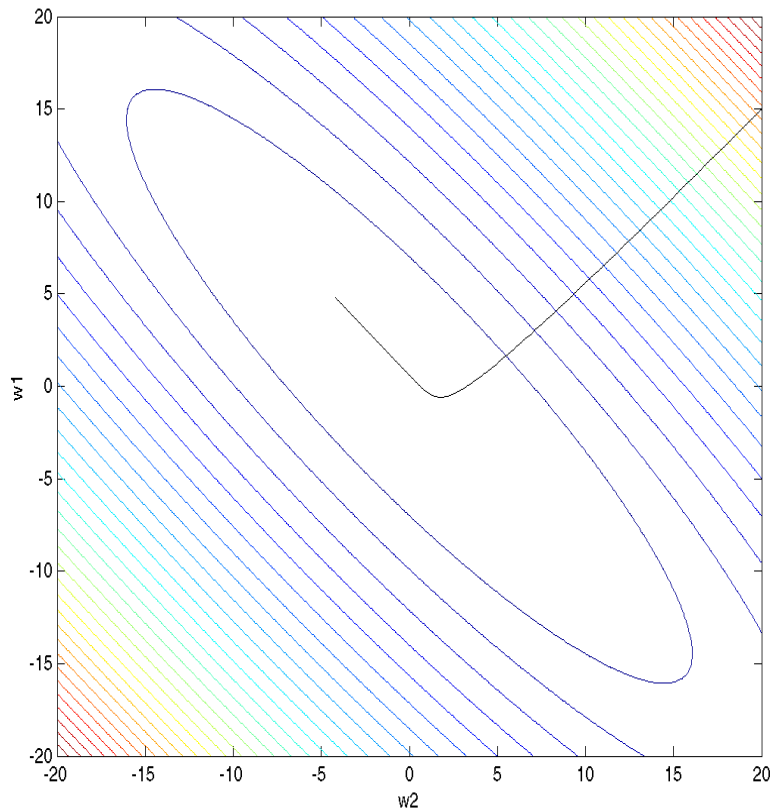
Joonis 22 kujutabki kiireima languse meetodi kaalude vektori trajektoori. Antud juhul on sisendsignaaliks valge müra, st sisendsignaali autokorrelatsioonimaatriks on ühikmaatriks $\mathbf{R} =$

$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Sellisel juhul on sisendsignaali korrelatsioonimaatriksi omaväärtused omavahel võrdsed $\lambda_1 = \lambda_2 = 1$ ning vea pinda iseloomustavad samakõrgusjooned on ringid. Algoritm koondub mööda sirgjoont vea pinna miinimumpunkti poole. Sisendsignaali ja soovitud signaali vaheline ristkorrelatsioonivektor on antud näites $\mathbf{p} = \begin{bmatrix} 1 \\ 0,1 \end{bmatrix}$. Vea pinna miinimumpunkt asub Wieneri lahendiga määratud koordinaatidel $\mathbf{w}_o = \mathbf{R}^{-1}\mathbf{p} = \begin{bmatrix} 1,0 \\ 0,1 \end{bmatrix}$.



Joonis 23. Kiireima languse meetodi kaalude vektori trajektoor korreleeritud sisendsignaali korral

Joonis 23 kujutab samuti kiireima languse meetodi kaalude vektori trajektoori vea pinnal, kusjuures sisendsignaali autokorrelatsioonimaatriks on antud juhul $\mathbf{R} = \begin{bmatrix} 1 & 0,5 \\ 0,5 & 1 \end{bmatrix}$, omaväärtustega $\lambda_1 = 0,5$ ja $\lambda_2 = 1,5$. Sisendsignaali ja soovitud signaali ristkorrelatsiooni vektor on endiselt $\mathbf{p} = \begin{bmatrix} 1 \\ 0,1 \end{bmatrix}$ ning Wieneri lahend, mis määrab vea pinna miinimumpunkti asukoha, on seega $\mathbf{w}_o = \mathbf{R}^{-1}\mathbf{p} = \begin{bmatrix} 1,27 \\ 0,53 \end{bmatrix}$. Nagu jooniselt näha, ei ole samakõrgusjooned enam ringid, vaid ellipsid. Samuti ei ole kaalude vektori trajektoor enam sirgjoon, nagu ta oli valge müra korral.



Joonis 24. Kiireima languse meetodi kaalude vektori trajektoor tugevalt korreleeritud sisendsignaali korral

Joonis 24 kujutab jällegi kiireima languse meetodi kaalude vektori trajektoori vea pinnal. Seekord on sisendsignaal tugevalt korreleeritud, autokorrelatsioonimaatriksiga $R = \begin{bmatrix} 1 & 0,9 \\ 0,9 & 1 \end{bmatrix}$, mille omaväärtused on $\lambda_1 = 0,1$ ja $\lambda_2 = 1,9$. Sisendsignaali ja soovitud signaali ristkorrelatsiooni vektor on taas $\mathbf{p} = \begin{bmatrix} 1 \\ 0,1 \end{bmatrix}$ ning Wieneri lahend, mis määrab vea pinna miinimumpunkti asukoha seega

$$\mathbf{w}_o = \mathbf{R}^{-1}\mathbf{p} = \begin{bmatrix} 4,7 \\ -4,2 \end{bmatrix}. \text{ Nagu jooniselt näha, on vea pinna samakõrgusjooned tugevalt väljavenitatud}$$

ellipsid ning kaalude vektori trajektoor on tugevalt kõverdunud.

5. Adaptiivsed LMS filtrid

Selles peatükis käsitleme adaptiivseid LMS filtreid. LMS on lühend ingliskeelsest Least Mean Square, mis tähendab vähimat ruutkeskmist. Tuleb kohe öelda, et sellist nime kannab LMS algoritm ajaloolistel põhjustel, vaatamata sellele, et ta tegelikult ei ole vähima ruutkeskmise vea mõttes optimaalne. LMS algoritmi pakkusid 1960. aastal välja Widrow ning Hoff [18]. Algoritm on tänu oma lihtsusele võitnud suure populaarsuse ning on (koos oma variatsioonidega) kindlasti praktikas enim kasutatud adaptiivne algoritm.

Algoritm kuulub nn stohhastilise gradiendi algoritmide hulka. Seda vastandina kiireima languse meetodile, kus on kasutusel deterministlik gradient. Kiirema languse meetodi puhul me eeldasime, et sisendsignaali korrelatsioonimaatriks ning sisendsignaali ja soovitud signaali vaheline korrelatsioonivektor on teada. Praktikas meil sageli sellist informatsiooni ei ole. Järgnevas vaatlemegi, kuidas konstrueerida lihtne ja praktiline adaptiivse filtreerimise algoritm ilma signaalide ja nende seoste kohta lisainformatsiooni nõudmata ning uurime ka tema omadusi.

Tuletame meelde, et kiireima languse meetodi puhul kasutasime gradienti

$$\nabla J(n) = -2\mathbf{p} + 2\mathbf{R}\mathbf{w}(n) \quad (5.1)$$

kus \mathbf{p} on üle ansambli keskmistatud sisendsignaali ja soovitud signaali vaheline korrelatsioonivektor ja \mathbf{R} samuti üle ansambli keskmistatud sisendsignaali korrelatsioonimaatriks. Kiireima languse meetodi rekursiooniks oli

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu[\mathbf{p} - \mathbf{R}\mathbf{w}(n)], \quad n = 0, 1, 2, \dots \quad (5.2)$$

Saamaks LMS algoritmi, asendame korrelatsioonid võrrandites (5.1) ja (5.2) keskmistamata hetkehinnangutega

$$\begin{aligned} \hat{\mathbf{R}}(n) &= \mathbf{x}(n)\mathbf{x}^H(n) \\ \hat{\mathbf{p}}(n) &= \mathbf{x}(n)d(n) \end{aligned} \quad (5.3)$$

saades gradiendi jaoks avaldise

$$\hat{\nabla} J(n) = -2\mathbf{x}(n)d(n) + 2\mathbf{x}(n)\mathbf{x}^H(n)\hat{\mathbf{w}}(n) \quad (5.4)$$

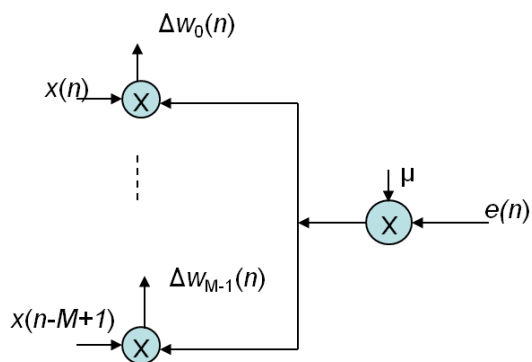
Selline gradiendi hinnang on juhuslik suurus ja sellest ka nimi – stohhastiline gradient. Kasutades äsjaleitud gradiendi hetkehinnangut tegeliku gradiendi asemel valemis (4.1), saame

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \mu\mathbf{x}(n)[d^*(n) - \mathbf{x}^H(n)\hat{\mathbf{w}}(n)], \quad n = 0, 1, 2, \dots \quad (5.5)$$

Kandilistes sulgudes olev tegur on võrdne veasignaali ajahetkel n ja seega saame LMS koefitsientide rekursiooni lõplikuks kujuks

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \mu\mathbf{x}(n)e^*(n), \quad n = 0, 1, 2, \dots \quad (5.6)$$

Joonis 25 kujutab valemit (5.6) graafiliselt.



Joonis 25. LMS algoritmi rekursioon

Seega võime LMS algoritmi esitada kolme võrrandina. Esiteks, filtreerimine

$$y(n) = \hat{\mathbf{w}}^H(n)\mathbf{x}(n) \tag{5.7}$$

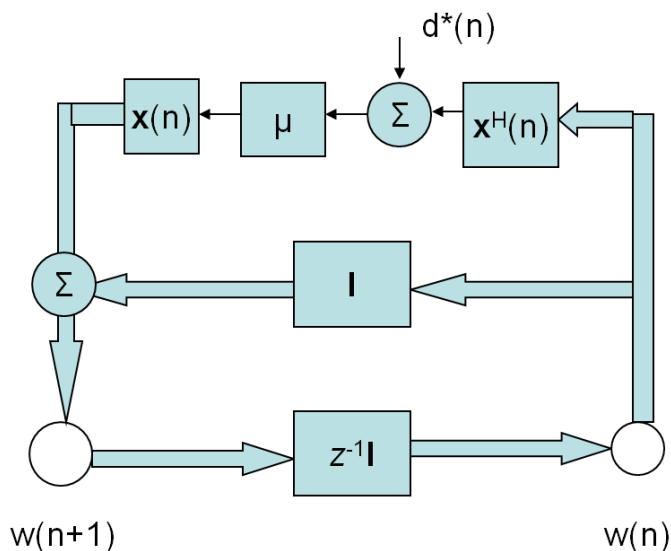
teiseks, veasignaali arvutus

$$e(n) = d(n) - y(n) \tag{5.8}$$

ning lõpuks koefitsientide uuendamine

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \mu \mathbf{x}(n)e^*(n) \tag{5.9}$$

Oluline on, et arvutused teostataks just nimelt sellises järjekorras.



Joonis 26. LMS algoritmi signaaligraaf

Joonis 26 kujutab LMS algoritmi signaaligraafina. Paneme tähele, et igal iteratsioonil on vajalikud ainult kõige viimane sisendvektor $\mathbf{x}(n)$ ja soovitud signaali diskreet $d(n)$. Peale selle vajab algoritm mälu ainult koefitsientide vektori $\hat{\mathbf{w}}$ jaoks.

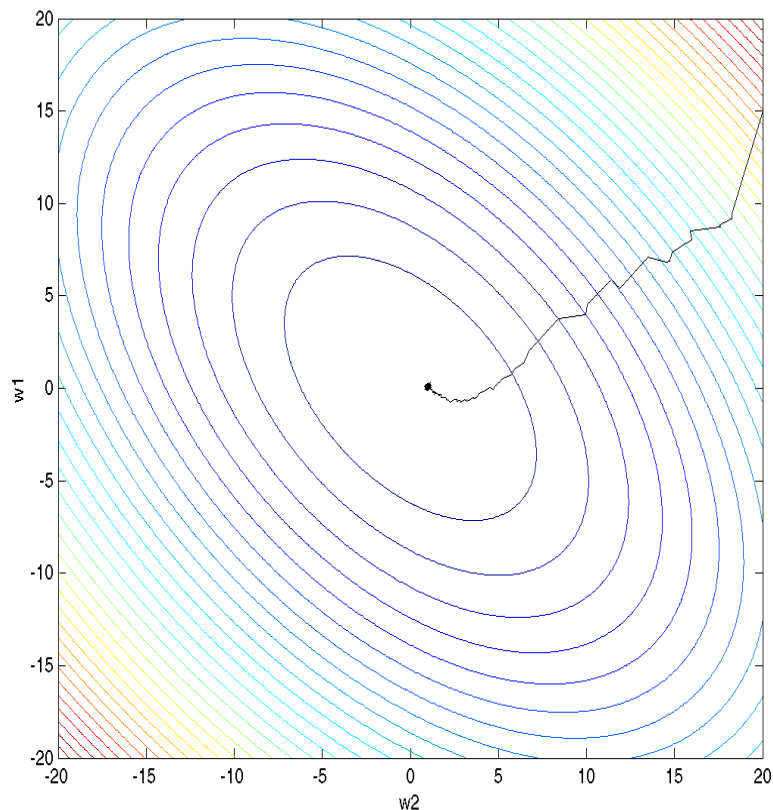
Nagu eelnevalt nägime, kasutab algoritm \mathbf{R} ja \mathbf{p} hetkehinnanguid, mis on suure dispersiooniga. See viib mõttele, et selliselt konstrueeritud algoritm ei saa hästi töötada. Õnneks tasakaalustab

keskmistamata hinnangute kasutamist algoritmi enese tagasisidest tulenev keskmistamisefekt ja sellest tulenevalt töötab algoritm aktsepteeritavalt, nagu me ka varsti näeme.

LMS algoritmi realiseerimiseks on vaja $2M+1$ kompleksset korrutust ja $2M$ kompleksset liitmist iteratsiooni kohta. Seega sõltub arvutuste hulk filtrikoefitsientide arvust lineaarselt ehk algoritmi keerukus on $O(M)$.

Näide

Vaatleme LMS algoritmi kaalude trajektoori vea pinnal.



Joonis 27. LMS algoritmi kaalude trajektoori vea pinnal

Joonis 27 kujutabki ühte konkreetset LMS kaalude trajektoori. Sisendsignaali korrelatsioonimaatriks on antud juhul $\mathbf{R} = \begin{bmatrix} 0,9 & 0,45 \\ 0,45 & 0,9 \end{bmatrix}$, omaväärtustega $\lambda_1 = 0,46$ ja $\lambda_2 = 1,35$. Sisendsignaali ja soovitud signaali vaheline ristkorrelatsiooni vektor on $\mathbf{p} = \begin{bmatrix} 0,95 \\ 0,45 \end{bmatrix}$. Vea pinna miinimumi asukoht on seega

$$\mathbf{w}_o = \mathbf{R}^{-1}\mathbf{p} = \begin{bmatrix} 1,06 \\ -0,03 \end{bmatrix}. \text{ Näites on kasutatud LMS algoritmi, sammu pikkusega } \mu = 0,05.$$

Nagu jooniselt nähtub, koondub LMS algoritm mööda teed, mis on küll üldjoontes sarnane kiireima languse meetodi koondumisteega (vt Joonis 22 kuni Joonis 24), kuid ei ole sama sujuv.

LMS algoritmi analüüs

Vaatamata LMS algoritmi enda lihtsusele, osutub tema analüüs ootamatult keeruliseks ning on aktiivse teadustöö objektiks tänaseni. On ju tegu süsteemiga, mis on ajas muutuv, juhuslik ning kaalude uuendamise võrrandist (5.9) tulenevalt ka mittelineaarne. Need on keerukused, mis ka eraldi võetuina raskendaksid süsteemi analüüsi, kuid antud juhul esinevad kõik kolm üheaegselt. Käesolevas peatükis esitame ühe algoritmi klassikalistest analüüsides. See analüüs põhineb teatud lihtsustavatel eeldustel, mida kutsutakse sõltumatuse teooriaks. Sõltumatuse teooria eeldused on järgmised:

1. Sisendsignaali vektorid $\mathbf{x}(i)$, $i=1, \dots, n$ on üksteisest statistiliselt sõltumatud.
2. Ajahetkel n on sisendvektor $\mathbf{x}(n)$ statistiliselt sõltumatu kõigist eelnevatest soovitud signaali diskreetidest $d(i)$, $i=1, \dots, n-1$.
3. Ajahetkel n on soovitud signaal $d(n)$ statistiliselt sõltumatu kõigist soovitud signaali eelnevatest diskreetidest $d(i)$, $i=1, \dots, n-1$.
4. Sisendsignaali vektorid $\mathbf{x}(i)$ ja soovitud signaal $d(i)$ on vastastikku normaaljaotusega iga i korral.

Eelnevalt (5.5) nägime, et LMS algoritmi koefitsientide vektor sõltub ainult sisendsignaali vektoritest $\mathbf{x}(n)$, $\mathbf{x}(n-1), \dots, \mathbf{x}(1)$, soovitud signaali diskreetidest $d(n)$, $d(n-1), \dots, d(1)$ ning koefitsientide vektori algolekust ajahetkel 0, $\hat{\mathbf{w}}(0)$. Tähistame koefitsientide vea vektori

$$\boldsymbol{\varepsilon}(n) = \hat{\mathbf{w}}(n) - \mathbf{w}_o \quad (5.10)$$

kus \mathbf{w}_o on optimaalne koefitsientide vektor (Wieneri lahend). Uurides LMS võrrandeid ning esimest ja teist sõltumatuse teooria eeldust, jõuame järeldusele, et kaalude vektor ajahetkel n , $\hat{\mathbf{w}}(n)$ ning seega ka kaalude vigade vektor $\boldsymbol{\varepsilon}(n) = \hat{\mathbf{w}}(n) - \mathbf{w}_o$, on statistiliselt sõltumatud nii ajahetkel n saabuvast uuest sisendvektorist $\mathbf{x}(n)$ kui ka uuest soovitud signaali diskreedist $d(n)$.

LMS kaalude uuendamise võrrandist (5.9) saame

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \mu \mathbf{x}(n) [d^*(n) - \mathbf{x}^H(n) \hat{\mathbf{w}}(n)] \quad (5.11)$$

Asendus valemist (5.10) annab

$$\hat{\mathbf{w}}(n+1) - \mathbf{w}_o = \hat{\mathbf{w}}(n) - \mathbf{w}_o + \mu \mathbf{x}(n) d^*(n) - \mu \mathbf{x}(n) \mathbf{x}^H(n) (\boldsymbol{\varepsilon}(n) + \mathbf{w}_o) \quad (5.12)$$

Seega saame koefitsientide vea vektori sammult n sammule $n+1$ uuendamise võrrandiks

$$\boldsymbol{\varepsilon}(n+1) = \boldsymbol{\varepsilon}(n) - \mu \mathbf{x}(n) \mathbf{x}^H(n) \boldsymbol{\varepsilon}(n) + \mu \mathbf{x}(n) (d^*(n) - \mathbf{x}^H(n) \mathbf{w}_o) \quad (5.13)$$

Paneme tähele, et viimastes sulgudes olev avaldis on võrdne optimaalse Wieneri filtri veasignaali ajahetkel n . Lihtsustades nüüd eelnevat võrrandit, saame esitada koefitsientide vea vektori uuendamise võrrandi mugavamal kujul

$$\boldsymbol{\varepsilon}(n+1) = (\mathbf{I} - \mu \mathbf{x}(n) \mathbf{x}^H(n)) \boldsymbol{\varepsilon}(n) + \mu \mathbf{x}(n) e_o^*(n) \quad (5.14)$$

Edasi vaatleme veasignaali dispersiooni. Veasignaali saab esitada koefitsientide vea vektori kaudu kujul

$$e(n) = d(n) - \hat{\mathbf{w}}^H(n)\mathbf{x}(n) = d(n) - (\mathbf{w}_o + \boldsymbol{\varepsilon}(n))^H \mathbf{x}(n) = e_o(n) - \boldsymbol{\varepsilon}^H(n)\mathbf{x}(n) \quad (5.15)$$

Seega avaldub veasignaali dispersioon kui

$$\begin{aligned} J(n) &= E\left[|e(n)|^2\right] = E\left[(e_o(n) - \boldsymbol{\varepsilon}^H \mathbf{x}(n))(e_o^*(n) - \mathbf{x}(n)^H \boldsymbol{\varepsilon})\right] = \\ &= E[e_o(n)e_o^*(n)] + 2\operatorname{Re}\{E[e_o^*(n)\boldsymbol{\varepsilon}^H \mathbf{x}(n)]\} + E[\boldsymbol{\varepsilon}^H \mathbf{x}(n)\mathbf{x}(n)^H \boldsymbol{\varepsilon}] \end{aligned} \quad (5.16)$$

Saadud valemi esimene liidetav on optimaalse Wieneri filtri veasignaali dispersioon, seega vähim võimalik veasignaali võimsus. Rakendame nüüd ortogonaalsuse printsiipi, mis ütleb, et optimaalse Wieneri filtri veasignaali ning sisendsignaali on ortogonaalsed. Seega on eelnevas valemis teine liidetav nulline. Järelikult võime me eelneva valemi kirjutada kujul

$$J(n) = J_{\min} + E[\boldsymbol{\varepsilon}^H \mathbf{x}(n)\mathbf{x}(n)^H \boldsymbol{\varepsilon}] = J_{\min} + \operatorname{Tr}\{E[\mathbf{x}(n)\mathbf{x}(n)^H]E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^H]\} \quad (5.17)$$

Kus Tr tähistab maatriksi jälje arvutamise operatsiooni. Maatriksi jälg on lineaaralgebrast tuntud operaator ja ta on võrdne maatriksi peadiagonaali elementide summaga. Eelnevas väites on kasutatud fakti, et skalaarist arvatud maatriksi jälg on võrdne skalaari endaga ning maatriksi jälje omadust

$$\operatorname{Tr}\{\mathbf{ABC}\} = \operatorname{Tr}\{\mathbf{BCA}\} = \operatorname{Tr}\{\mathbf{CAB}\} \quad (5.18)$$

Esimene tegur valemi (5.17) maatriksi jälje operaatori argumendis on võrdne sisendsignaali korrelatsioonimaatriksiga \mathbf{R} ning teine tegur on koefitsientide vigade korrelatsioonimaatriks, mille me tähistame sümboliga \mathbf{K} saades

$$J(n) = J_{\min} + \operatorname{Tr}[\mathbf{RK}(n)] \quad (5.19)$$

Kaalude vigade korrelatsioonimaatriksi uurimiseks kasutame nn otsese keskmistamise meetodit [5], mille kohaselt võime kaalude vea vektori uuendamise võrrandis (5.14) asendada esimese teguri tema keskmistatud variandiga

$$E[\mathbf{I} - \mu\boldsymbol{\alpha}(n)\mathbf{x}^H(n)] = \mathbf{I} - \mu\mathbf{R} \quad (5.20)$$

saades

$$\boldsymbol{\varepsilon}(n+1) \approx (\mathbf{I} - \mu\mathbf{R})\boldsymbol{\varepsilon}(n) + \mu\boldsymbol{\alpha}(n)e_o^*(n) \quad (5.21)$$

Kasutades eeltoodud avaldist kaalude vigade korrelatsioonimaatriksi leidmiseks, saame koefitsientide vea korrelatsioonimaatriksi rekursiooniks

$$\mathbf{K}(n+1) = E[\boldsymbol{\varepsilon}(n+1)\boldsymbol{\varepsilon}^H(n+1)] = (\mathbf{I} - \mu\mathbf{R})\mathbf{K}(n)(\mathbf{I} - \mu\mathbf{R}) + E[\mu^2 e_o^*(n)\mathbf{x}^H(n)e_o(n)] \quad (5.22)$$

Eelnevast oleme elimineerinud ristliikmed tüüpi $E[\mu e_o(n)(\mathbf{I} - \mu\mathbf{R})\boldsymbol{\varepsilon}(n)\mathbf{x}^H(n)]$, kuna nad on sõltumatuse teooriast tuleneva vektorite $\boldsymbol{\varepsilon}(n)$ ning $\mathbf{x}(n)$ statistilise sõltumatuse põhjal nullised. Viimane liidetav on keskvärtus nelja juhusliku suuruse korrutisest ja tema lihtsustamiseks kasutame matemaatikast tuntud Gaussi faktoriseerimise teoreemi, mis ütleb, et kui kompleksed juhuslikud suurused x_1, x_2, x_3 ja x_4 on vastastikku normaaljaotusega, kehtib

$$E[x_1 x_2^* x_3 x_4^*] = E[x_1 x_2^*]E[x_3 x_4^*] + E[x_1 x_4^*]E[x_3 x_2^*] \quad (5.23)$$

Kasutades ülaltoodut, saame kaalude vigade korrelatsioonimaatriksi rekursiooniks

$$\mathbf{K}(n+1) = (\mathbf{I} - \mu\mathbf{R})\mathbf{K}(n)(\mathbf{I} - \mu\mathbf{R}) + \mu^2 J_{\min} \mathbf{R} \quad (5.24)$$

Siin on oluline tähele panna, et eeltoodud rekursiooni teine liige ei sõltu ajast ning seega ei koonu LMS algoritmi iseloomustav kaalude vigade korrelatsioonimaatriks nulliks. Valemist (5.19) järeldub nüüd, et LMS algoritmi viga jääb suuremaks Wieneri filtri veast nn lisavea dispersiooniga

$$J_{lisa}(n) = J(n) - J_{\min} = Tr[\mathbf{R}\mathbf{K}(n)] \quad (5.25)$$

võrra.

Rusikareeglid

Toome siin ära mõned lihtsad ning praktikas kasutamiseks mugavad reeglid, ilma neid tuletamata.

LMS algoritm koondub ruutkeskmise vea mõttes, kui tema sammu suurus rahuldab võrratust

$$0 < \mu < \frac{2}{\lambda_{\max}} \quad (5.26)$$

kus λ_{\max} on sisendsignaali autokorrelatsioonimaatriksi maksimaalne omaväärtus.

Seega on LMS algoritmi koonduvuse tingimus sama, mis kiireima languse meetodi puhulgi.

LMS algoritmi lisaviga väljakujunenud režiimis $J_{lisa}(\infty)$ on väiksem kui minimaalne ruutkeskmise viga, juhul kui algoritmi samm μ on valitud selliselt, et on rahuldatud võrratus

$$\sum_{i=1}^M \frac{2\lambda_i}{2 - \mu\lambda_i} < 1 \quad (5.27)$$

LMS algoritmi kõrvalehäälestus avaldub kujul

$$M = \frac{J_{lisa}(\infty)}{J_{\min}} = \sum_{i=1}^M \frac{\mu\lambda_i}{2 - \mu\lambda_i} \quad (5.28)$$

Nagu eelnevas veendusime, määrab selle, kas LMS algoritm koondub või mitte, sisendsignaali korrelatsioonimaatriksi suurim omaväärtus (5.26). Tavaliselt aga kas ei ole see teada või siis kujutaks korrelatsioonimaatriksi omaväärtuste leidmine olulist arvutusvõimsuse lisakulu, mida me meeleldi väldiksime. On aga teada, et kuna maatriksi jälg on avaldatav ka tema omaväärtuste summana, ei saa sisendsignaali korrelatsioonimaatriksi maksimaalne omaväärtus olla suurem tema jäljest $\lambda_{\max} \leq Tr(\mathbf{R})$, seega LMS algoritm koondub, kui on rahuldatud tingimus

$$0 < \mu < \frac{2}{Tr(\mathbf{R})} \quad (5.29)$$

Kuna \mathbf{R} on statsionaarse protsessi korrelatsioonimaatriks, on tal Toeplitzi struktuur ning tema peadiagonaali elemendid on kõik võrdsed sisendsignaali dispersiooniga. Sellest tulenevalt võime valemi (5.29) kirjutada kujul

$$0 < \mu < \frac{2}{\sum_{k=0}^{M-1} E[x(n-k)]^2} = \frac{2}{M\sigma_x^2} \quad (5.30)$$

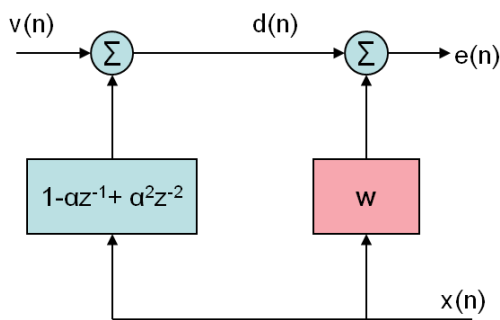
Väikese sammu suuruse korral on kõrvalehäälestus ligikaudselt arvutatav kui

$$M = \frac{J_{ex}(\infty)}{J_{\min}} = \sum_{i=1}^M \frac{\mu\lambda_i}{2 - \mu\lambda_i} \approx \frac{\mu}{2} \sum_{i=1}^M \lambda_i = \frac{\mu}{2} \sum_{i=1}^M E[x(n-k)]^2 \quad (5.31)$$

Seega on algoritmi kõrvalehäälestus seda suurem, mida pikem on algoritmi poolt kasutatav filter. See on ka loomulik tulemus, kuna pikema filtri puhul on vaja hinnata rohkem tundmatuid koefitsiente ja see annab ka suurema koguvea.

Näide

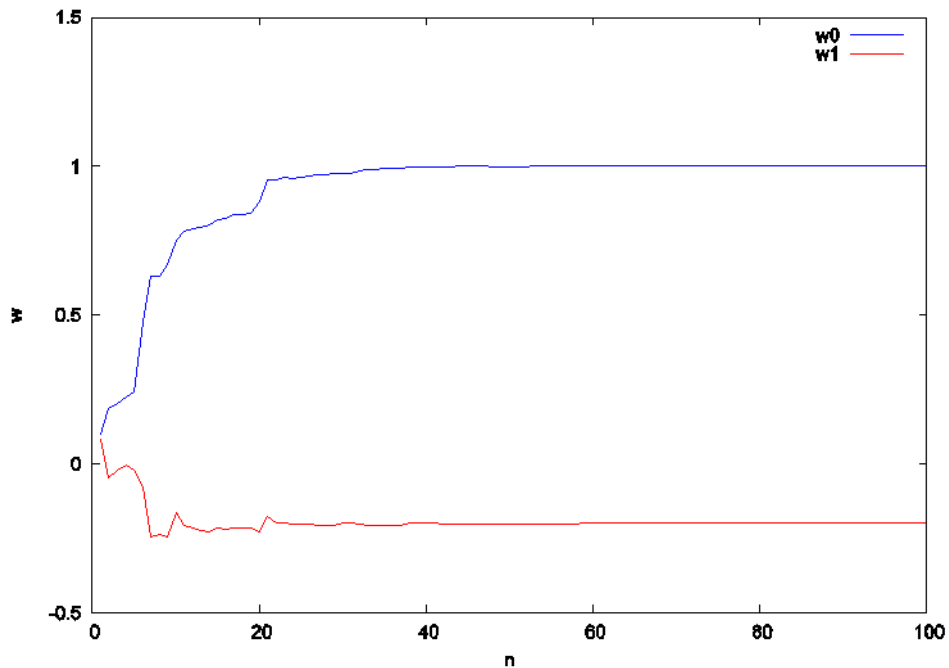
Joonis 28 kujutab adaptiivset filtrit, mis on ühendatud süsteemi identifitseerimise ülesande lahendamiseks. Identifitseeritav süsteem on lõpliku impulsskarakteristikuga filter, ülekandefunktsiooniga $1 - az^{-1} + a^2z^{-2}$, mille identifitseerimiseks kasutame adaptiivset filtrit koefitsientide vektoriga \mathbf{w} . Olgu $a = 0,2$. Sisendsignaalsiks $x(n)$ on valge müra võimsusega 1. Identifitseeritava süsteemi väljundile lisandub teie valge müra $v(n)$, võimsusega 10^{-6} . Protsessid x ja v on statistiliselt sõltumatud nullise keskväärtusega juhuslikud protsessid, mis alluvad normaaljaotusele. Kasutame LMS algoritmi sammu suurusega $\mu = 0,1$ ja olgu meie adaptiivses filtris \mathbf{w} , $M = 4$ koefitsienti algväärtustega $\mathbf{w}(0) = \mathbf{0}$.



Joonis 28. Näites uuritav süsteem

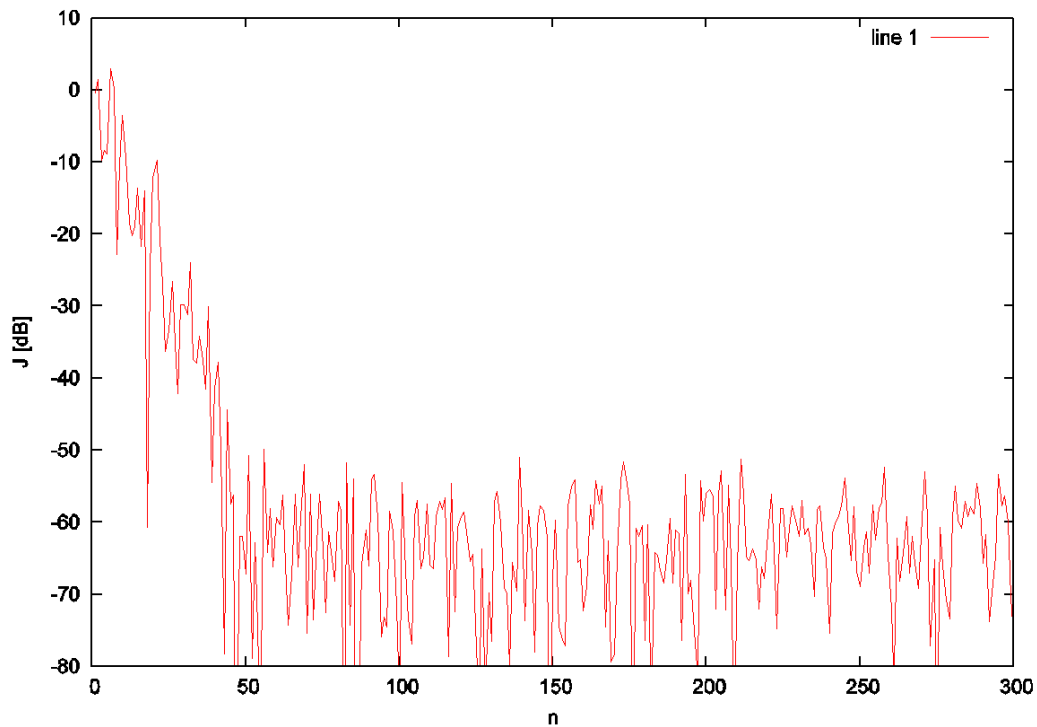
Joonis 29 näitab kahe esimese kaalu trajektoore aja funktsioonina. Näeme, et kaalud lähenevad otsitavatele väärtustele 1 ja 0.2 ning mõne aja möödudes stabiliseeruvad otsitavate väärtuste läheduses. Läheneviprotsess ei ole aga ühtlane ning koondumise käigus kaugenevad kaalud kohati otsitavatest väärtustest, et siis neile taas läheneda. Selline käitumine on LMS algoritmile tüüpiline, tuletame meelde, et LMS algoritm kasutab sisendvektori keskmistamata hetkväärtusi igal sammul uute kaalude leidmiseks. Joonise väiksuse tõttu tundub, nagu stabiliseeruksid kaalud mõne aja

möödudes täielikult, muutudes võrdseteks otsitavate väärtustega. Tegelikult seda aga ei juhtu ning kaalud jäävad juhuslikult, väikese amplituudiga, võnkuma otsitavate väärtuste ümber.



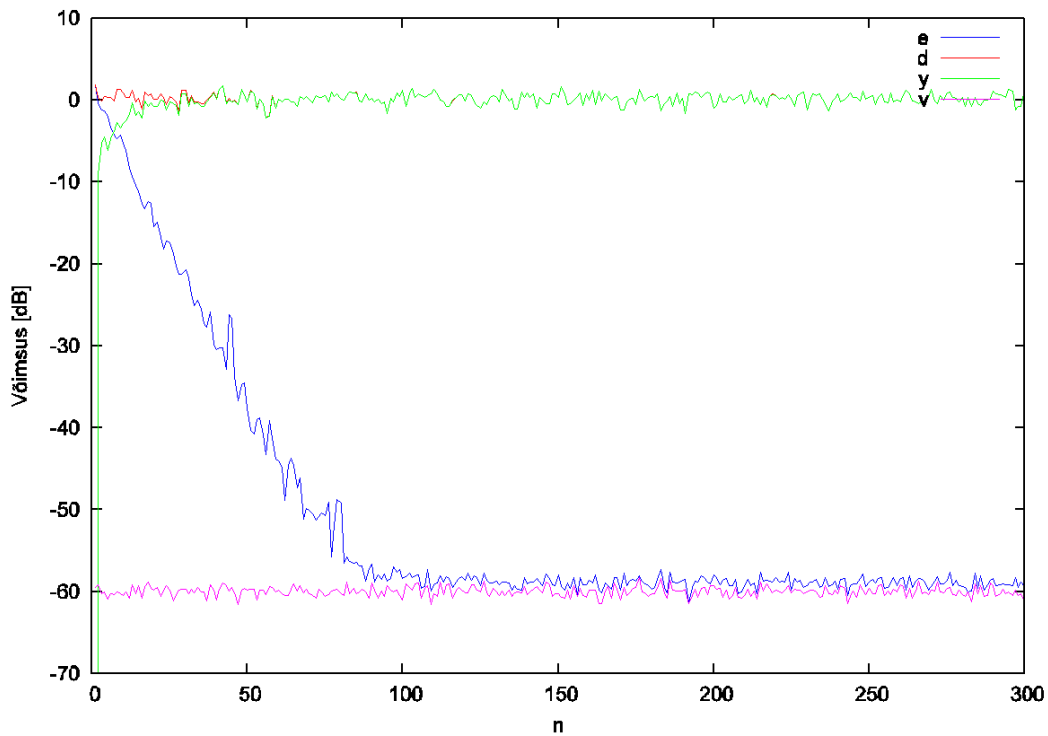
Joonis 29. Kaalude trajektoovid

Joonis 30 kujutab süsteemi identifitseerimise vea aja funktsioonina, esitatuna logaritmilisel skaalal (J on antud detsibellides). Näeme, et algselt viga väheneb ja jääb mõne aja möödudes võnkuma teatud taseme ümber, nagu ennustatud valemi (5.19) poolt



Joonis 30. Veasignaali ruut aja funktsioonina

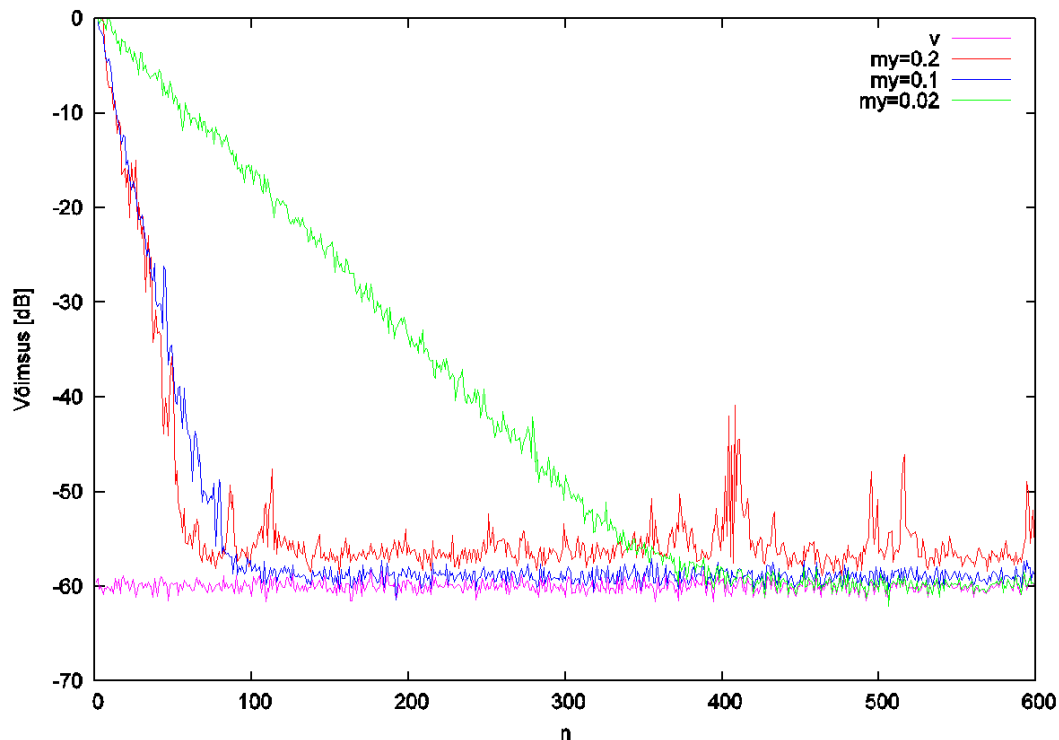
Joonis 31 kujutab algoritmiga seotud signaalide e , d , y ning v võimsuste arengut ajas. Nagu eelnevalt jooniselt nägime, pulseerivad katses esinevad signaalid (antud juhul veasignaali) juhuslikkusest tulenevalt märkimisväärses ulatuses, mis raskendab seaduspärasuste jälgimist. Vähendamaks signaalide juhuslikkusest tulenevat pulseerimist, esitame edaspidi ühes katses leitud signaali ruudu asemel 100 sõltumatu katse käigus leitud signaalide ruutude aritmeetilised keskmised. Teiste sõnadega, esitatavad graafikud on keskmistatud üle 100 realisatsiooni.



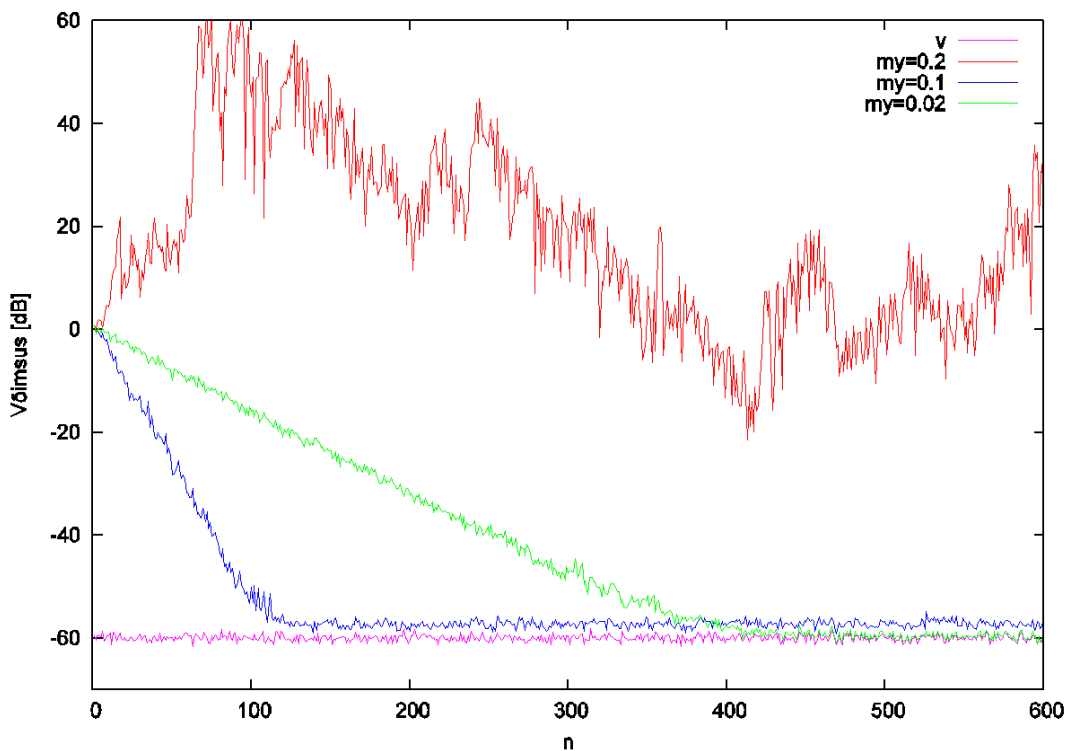
Joonis 31. Algoritmiga seotud signaalide ajaline areng

Nagu jooniselt näha, läheneb filtri väljundsignaal y soovitud signaalile d ja veasignaali e väheneb müra v taseme lähedale, jäädes sellest siiski pisut suuremaks.

Joonis 32 esitab LMS algoritmi veasignaali dispersiooni aja funktsioonina (õppekõvera) kolme erineva sammu väärtuse korral: $\mu = 0,2$, $\mu = 0,1$ ja $\mu = 0,02$. Lisaks on võrdluseks toodud mõõtemüra v dispersioon. Toodud kõverad on, nagu eelmiselgi joonisel, keskmistatud üle 100 sõltumatu realisatsiooni. Jooniselt võime näha, et mida väiksem on sammu suurus, seda aeglasemalt LMS algoritm koondub. Samas on hinnangu viga väljakujunenud režiimis seda väiksem, mida väiksem on sammu suurus. Seega tuleb LMS algoritmi puhul valida, mis on olulisem, kas kiire koondumus või väike lõppviga.



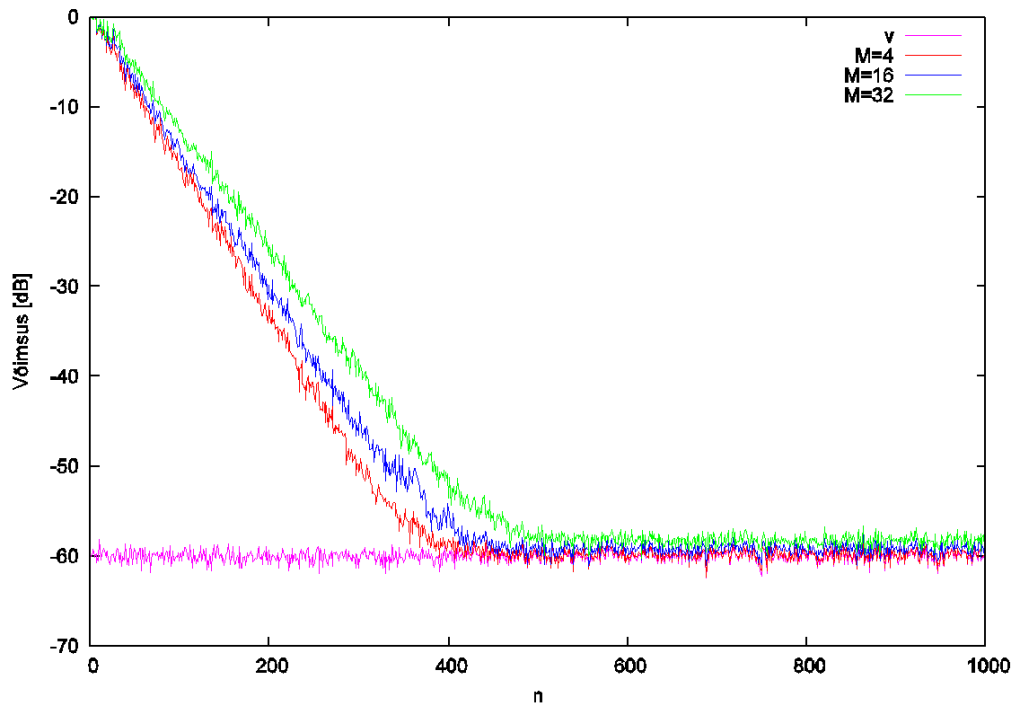
Joonis 32. Koonduvuse sõltuvus sammu suurusest



Joonis 33. Suurem mudel

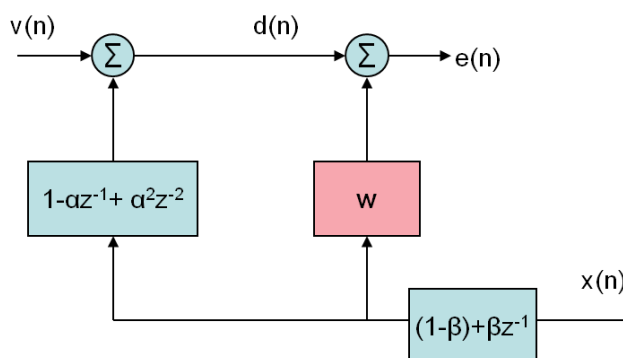
Joonis 33 kujutab taas LMS õppimiskõveraid, nagu ka Joonis 32. Situatsioon on täpselt sama nagu eelmises katses, ainsa vahega, et oleme kasutanud $M = 8$ kaaluga adaptiivset filtrit endise nelja kaalu asemel. Joonisel on näha, et algoritm koondub endiselt sammu suuruste $\mu = 0,1$ ja $\mu = 0,02$ korral. Kõige suurema sammu $\mu = 0,2$ korral aga koonduvust ei esine ning algoritmi lõppviga jääb ligikaudu

algvea tasemele. See on tingitud sellest, et pikem filter nõuab koondumiseks vastavalt valemile (5.30) väiksemat sammu suurust.



Joonis 34. Õppimiskõverate sõltuvus mudeli pikkusest

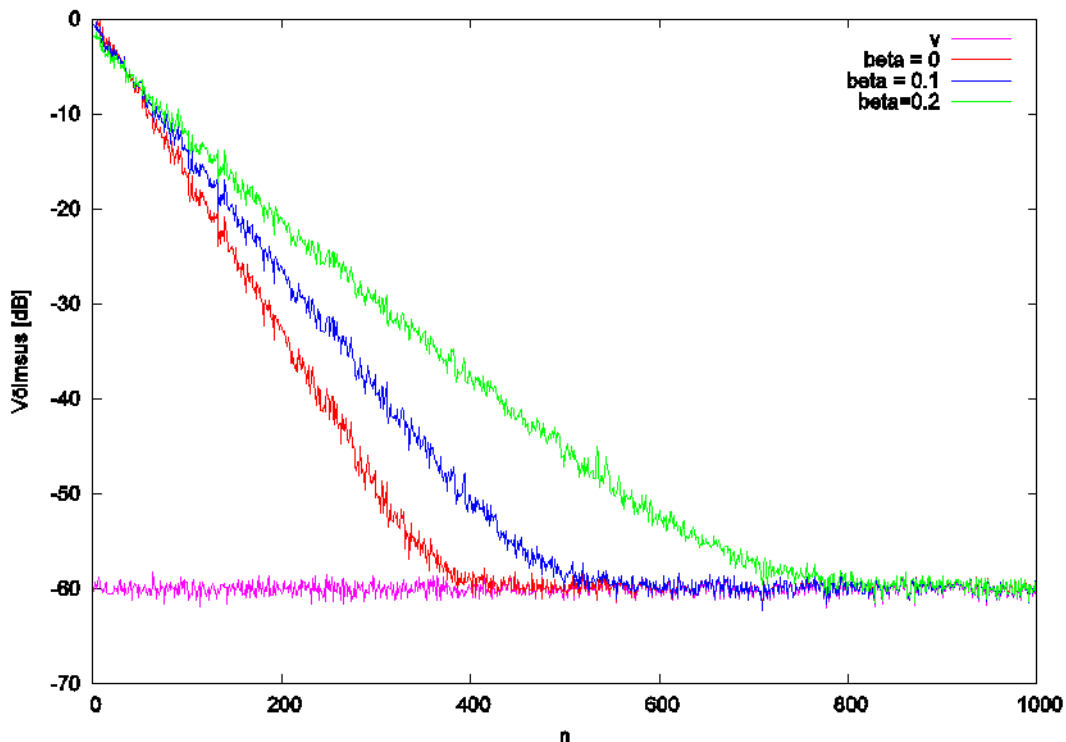
Joonis 34 näitab LMS õppimiskõverate sõltuvust valitud mudeli suurusest M juhtude $M = 4$, $M = 16$ ja $M = 32$ jaoks. LMS sammu suurus on kõigi kõverate arvutamisel sama $\mu = 0.02$. Näeme, et mida suurem on kasutatava mudeli pikkus, seda aeglasemalt koondub ka algoritmi ning seda suurem on algoritmi lõppviga.



Joonis 35. Uuritav süsteem värvilise sisendsignaali

Kui seni on meie katses sisendsignaal $x(n)$ olnud valge müra, siis nüüd lisame sisendsignaali spektri modifitseerimiseks filtri ning jälgime, kuidas LMS algoritmi omadused muutuvad, kui sisendsignaal on värviline. Värviline tähendab antud kontekstis signaali spektraaltiheduse erinevust konstandist.

Joonis 35 näitabki uuritavat süsteemi, kus sisendsignaali on filtreeritud eksponentsiaalselt keskmistava madalpääsfiltriga, mille ülekandefunktsioon on $(1 - \beta) + \beta z^{-1}$.

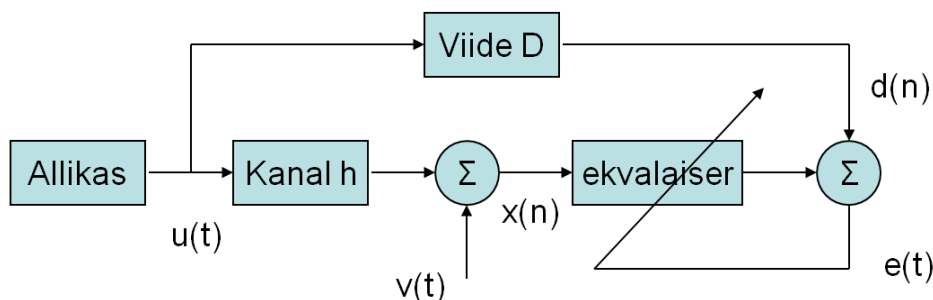


Joonis 36. Õppimiskõverate sõltuvus sisendsignaali korreleeritusest

Joonis 36 näitab LMS algoritmi õppimiskõveraid sõltuvana sisendsignaali rakendatud filtri koefitsiendist β . Mudeli suurus $M = 4$ ja sammu suurus $\mu = 0,02$. Nagu jooniselt näha, on LMS algoritmi koondumus seda aeglasem, mida suurem on β ehk mida korreleeritum on algoritmi poolt nähtav sisendprotsess. Seostades seda tulemust eelnevalt esitatud teooriaga, võime tõdeda, et juhul kui $\beta = 0$, on tegemist valge müraga, mille korrelatsioonimaatriksi omaväärtused on kõik võrdsed. β kasvades kasvab ka erinevus sisendsignaali korrelatsioonimaatriksi omaväärtuste vahel ning LMS algoritmi koondumus aeglustub. Sarnast fenomeni võisime jälgida ka kiireima languse meetodi puhul.

Näide. Adaptiivne sidekanali ekvalaiser

Ideaaljuhul ei too sidekanal edastatavasse signaali moonutusi. See on nii ainult juhul, kui sidekanali impulsskarakteristik on ühikimpulss ning sageduskarakteristik vastavalt konstantne. Praktikas see tavaliselt nii ei ole, kanali impulsskarakteristik erineb ühikimpulsist ning praktilise sidekanali läbinud signaalis esineb nähtus nimega sümbolitevaheline interferents. See tähendab, et impulsid, mida sümbolite edastamiseks kasutatakse, „määrduvad“ ajas laiali nii, et osa sümboli energiast saabub vastuvõtjasse alles samaaegselt järgnevate sümbolitega. Selliselt laialivalgunud signaal muutub häireks järgnevate sümbolite detekteerimisel, raskendades oluliselt detektori tööd ja suurendades detekteerimisvea tõenäosust. Probleemi lahendamiseks on kasutusel vastuvõtjasse paigutatud ekvalaiserid, mille ülesandeks on sidekanali sageduskarakteristiku ühtlustamine. Sidekanali sageduskarakteristik ei ole sidesüsteemi konstrueerimise ajal tavaliselt teada. Seetõttu peab ekvalaiser kohanduma olukorraga pärast töö alustamist – olema adaptiivne.



Joonis 37. Adaptiivne ekvalaiser

Joonis 37 kujutab sidesüsteemi, mis kasutab adaptiivset ekvalaiserit. Eeldame, et allika väljundsignaali diskreedid on üksteisest statistiliselt sõltumatud. Allikast väljuv signaal $u(t)$ läbib oma teel vastuvõtjasse sidekanali impulsskarakteristikuga $h(t)$ ning talle lisanduvad kanali mürad $v(t)$. Siin kasutatud sidekanali impulsskarakteristik sisaldab ka saatja ja vastuvõtja filtritest ning vastuvõtja filtrile järgnevast diskreetijast tulenevaid efekte. Eeldame, et kanali müra on valge, ning seega on müra erinevatel ajahetkedel võetud diskreedid üksteisega mittekorreleeritud. Ekvalaiseri ülesandeks on jälgida signaali kanali väljundis ning selle põhjal konstrueerida filter, mille järjestiklülitus kanaliga omaks ühikimpulsiga võrdset impulsskarakteristikut. Selle saavutamiseks kasutatakse soovitud signaalina $d(n)$ hilistatud allika väljundsignaali $u(n-D)$. Seega on adaptiivse filtri ülesanne formeerida kaalud, mille kasutamisel oleks adaptiivse filtri väljundsignaal hilistunud allika väljundsignaaliga nii sarnane kui võimalik. Hilistumine on siin vajalik sellepärast, et signaal hilistub nii sidekanalis kui ka ekvalaiseris ja me soovime saada ülesandele kausaalset lahendit.

Olgu meie allikas ja kanal iseloomustatavad võrranditega

$$u(n) = \pm 1$$

$$h(n) = \begin{cases} \frac{1}{2} \left[1 + \cos\left(\frac{2\pi}{W}(n-1)\right) \right], & n = 0, 1, 2 \\ 0, & \text{muudel juhtudel} \end{cases} \quad (5.32)$$

Siis saame ekvalaiseri sisendsignaali ning soovitud signaali avaldada kui

$$\begin{aligned} x(n) &= h(0)u(n) + h(1)u(n-1) + h(2)u(n-2) + v(n) \\ d(n) &= u(n-D) \end{aligned} \quad (5.33)$$

Kuna allika väljundsignaali diskreedid on erinevatel ajahetkedel üksteisega mittekorreleeritud ja kanali müra on valge, saame sisendsignaali korrelatsioonimaatriksi elementideks

$$\begin{aligned} r(0) &= h^2(0) + h^2(1) + h^2(2) + \sigma_v^2 \\ r(1) &= h(0)h(1) + h(1)h(2) \\ r(2) &= h(0)h(2) \end{aligned} \quad (5.34)$$

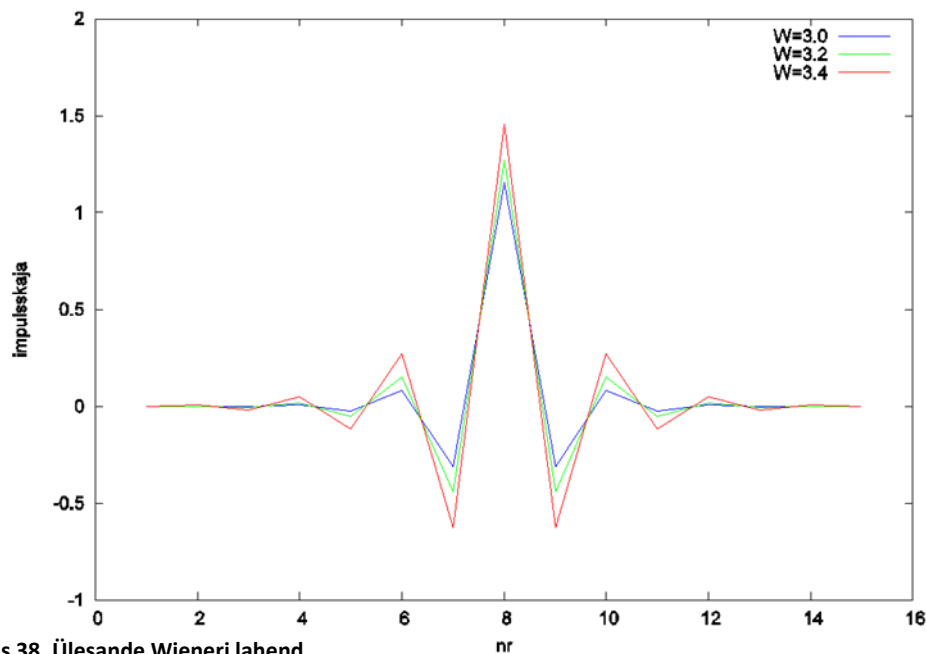
Sisendsignaali korrelatsioonimaatriksiks saame seega maatriksi, millel on nullist erinevad elemendid ainult peadiagonaalil ning kummalgi pool kahel kõrvaldiagonaalil.

$$\mathbf{R} = \begin{bmatrix} r(0) & r(1) & r(2) & 0 & \dots \\ r(1) & r(0) & r(1) & r(2) & \dots \\ r(2) & r(1) & r(0) & r(1) & \dots \\ 0 & r(2) & r(1) & r(0) & \dots \end{bmatrix} \quad (5.35)$$

Sisendsignaali ja soovitud signaali vastastikuse korrelatsiooni vektoriks tuleb

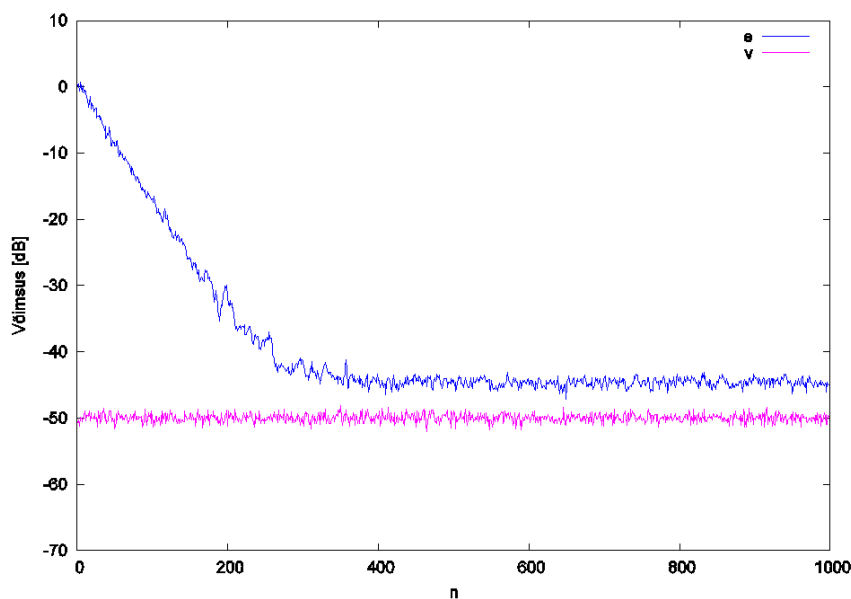
$$\mathbf{p} = [0 \quad \dots \quad h(0) \quad h(1) \quad h(2) \quad \dots \quad 0]^T \quad (5.36)$$

Valime kasutatava ekvalaiseri kaalude arvaks $M = 15$ ning viiteks $D = 8$. Siin tasub tähele panna, et me leiame lõpliku impulsskarakteristikuga filtri pöördfiltri, soovides lahendit, mis oleks samuti lõpliku impulsskarakteristikuga filter. Selline lahend on võimalik leida üldjuhul ainult ligikaudsena ning lahendi parendamiseks tuleks suurendada adaptiivse filtri pikkust M .



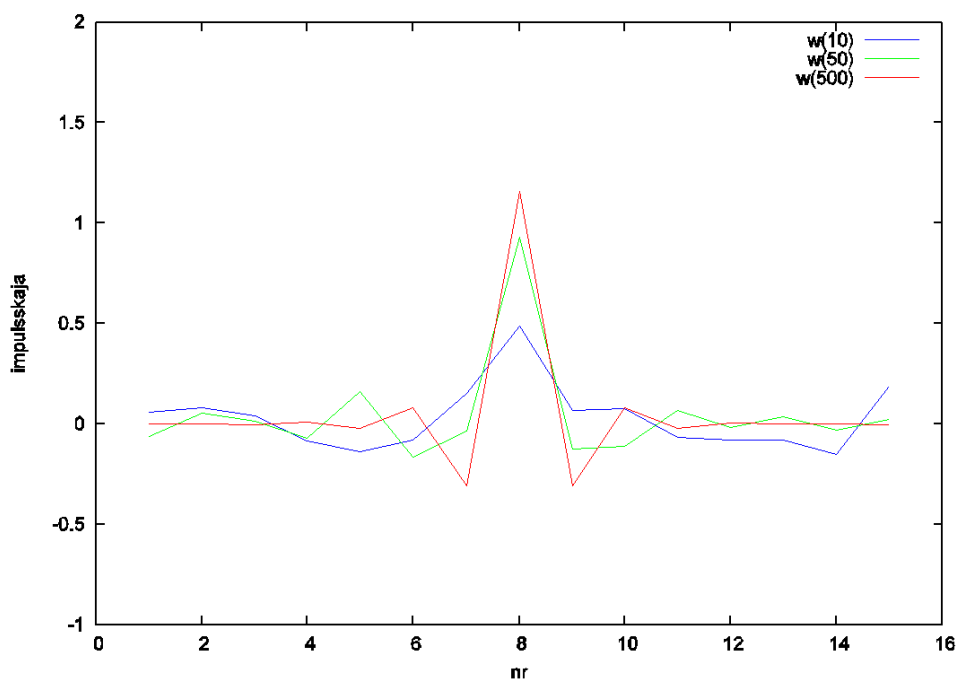
Joonis 38. Ülesande Wieneri lahend

Joonis 38 kujutab ülesande Wieneri lahendit $\mathbf{w} = \mathbf{R}^{-1}\mathbf{p}$ kolme erineva sidekanali impulsskarakteristiku valemis (5.32) esineva parameetri W korral. Paneme tähele, et mida suurem on W , seda suuremad on lahendi võnked.



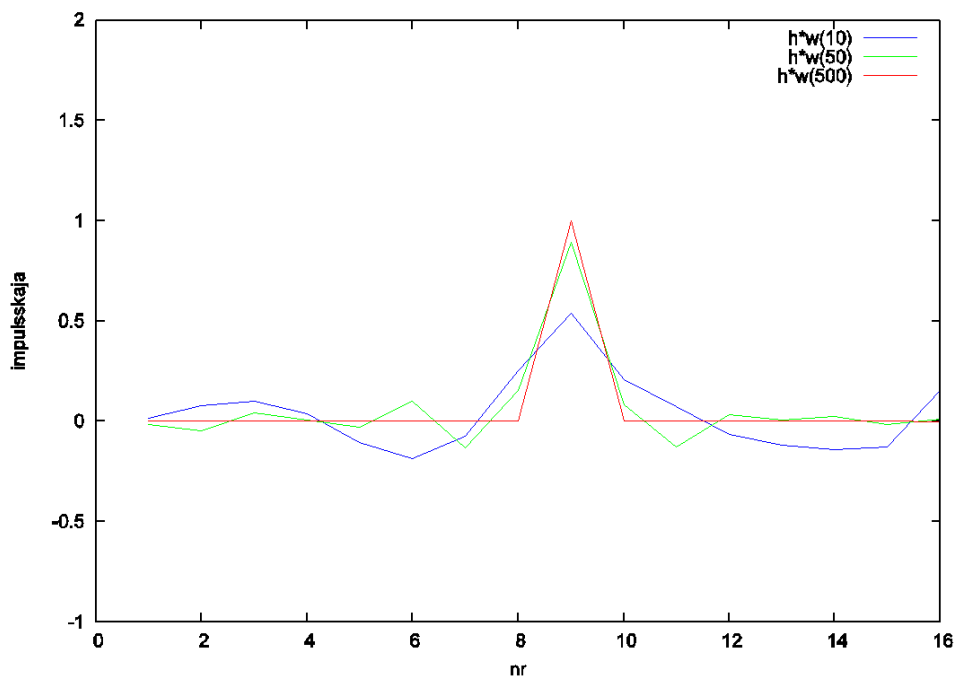
Joonis 39. Ekvalaiseri õppekõver

Joonis 39 kujutab adaptiivse ekvalaiseri õppekõverat, st veasignaali võimsuse sõltuvust ajast keskmistatuna üle 100 sõltumatu katse. Sisendsignaali võimsus on võrdne ühega. Võrdluseks on joonisel ära toodud ka kanali müra võimsus, milleks on 10^{-5} . Algoritmi sammu suuruseks on valitud $\mu = 0.05$, kaalude algolek on $\mathbf{w}(0) = \mathbf{0}$ ning kanalit iseloomustav parameeter W on valitud $W = 3$. Näeme, et algoritm koondub, kuid lõppviga jääb mürasignaali võimsusest kaugemale kui eelmises näites. See on tingitud antud ülesande Wieneri lahendi veast. Nimelt oleks kanali täpseks pööramiseks vaja lõppmatult pikka filtrit, meie aga kasutasime lõplikku filtrit pikkusega $M = 15$.



Joonis 40. Lahendi kujunemine adaptiooni käigus

Joonis 40 näitab ekvalaiseri kujunemist adaptatsiooni käigus, tuues ära impulsskarakteristikud peale kümnenadat, viiekümnenadat ning viiesajandat iteratsiooni kanali parameetriga $W = 3$ jaoks. Võrreldes antud joonisel kujutatut Wieneri lahendiga (Joonis 38), näeme, et esialgsed lahendid erinevad soovitud oluliselt, kuid viiesajandaks iteratsiooniks on algoritm koondunud ning saadud lahend on Wieneri lahendiga küllaltki sarnane. Kuna viimast asjaolu on silma järgi pisut raske täpselt hinnata, teeme lihtsa kontrolli. Tuletame meelde, et me soovisime kanali ja ekvalaiseri järjestikühenduse impulsskarakteristiku võimalikult suurt sarnasust ühikimpulsiga.



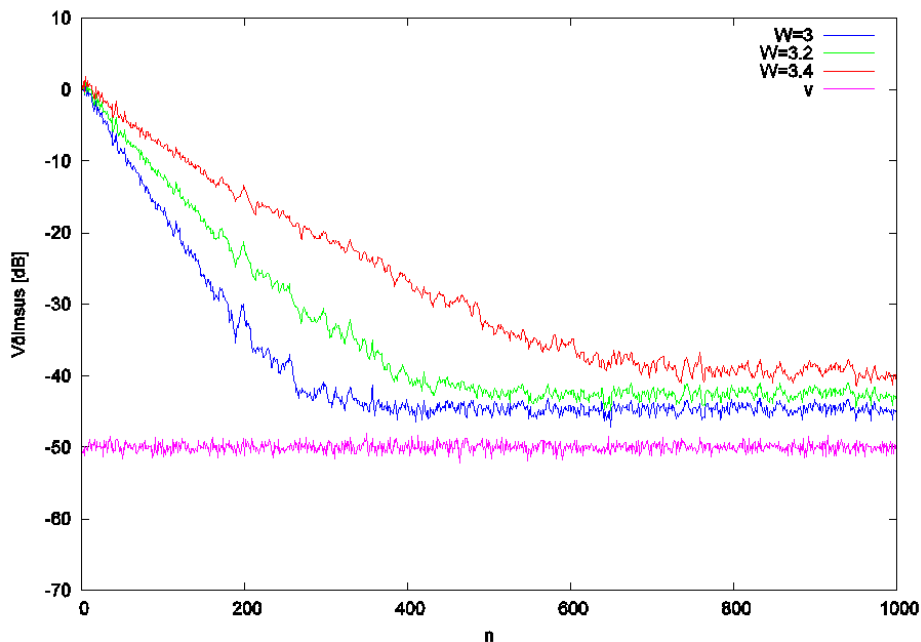
Joonis 41. Kanali ja ekvalaiseri järjestikühenduse impulsskarakteristik

Joonis 41 kujutabki kanali ja ekvalaiseri järjestikühenduse impulsskarakteristikut erinevates adaptatsioonistaadiumites: pärast kümnenadat, viiekümnenadat ning viiesajandat iteratsiooni. Tuleb tähele panna, et kuna meil on tegemist diskreetse ajaga, on olulised ainult täisarvulistele argumentidele vastavad funktsiooniväärtused. Näeme, et kujutatud impulsskarakteristik läheneb ühikimpulsile ning viiesajandaks iteratsiooniks on erinevus ühikimpulsist palja silmaga eristamatu.

Lõpuks vaatleme adaptatsiooni sõltuvust kanali impulsskarakteristiku parameetrist W ning selle kaudu sisendsignaali korrelatsioonimaatriksi omaväärtustest. Järgnevas tabelis on toodud korrelatsioonimaatriksi vähim ning suurim omaväärtus kolme kanali parameetri W väärtuse korral ning suurima ja vähima omaväärtuse suhe. Näeme, et W suurenedes suureneb omaväärtuste suhe kiiresti.

Tabel 1. Sisendsignaali korrelatsioonimaatriksi omaväärtused erinevate kanali parameetrite korral

W	3.0	3.2	3.4
λ_{\min}	0.260	0.156	0.081
λ_{\max}	2.22	2.58	2.94
$\lambda_{\max}/\lambda_{\min}$	8.54	16.5	35.1



Joonis 42. Koonduvuse sõltuvus kanali parameetrist

Joonis 42 kujutab üle saja realisatsiooni keskmistatud LMS õppekõveraid nende kolme kanali parameetri 3,0; 3,2 ja 3,4 korral, kusjuures muud tingimused on kõik samad. Tabel 1 näitas, et kanali parameetri kasvades suurenes kiiresti ka sisendsignaali korrelatsioonimaatriksi suurima ja vähima omaväärtuse suhe. Nagu jooniselt nähtub, kaasneb sellega adaptatsiooni märgatav aeglustumine.

Tähelepanelik lugeja on tõenäoliselt märganud, et siiani oleme eeldanud müra- ja moonutustevaba hilistunud allika väljundsignaali kättesaadavust vastuvõtjas, mis muudab kogu konstruktsiooni mõttetuks. Praktikas on skeem sellegipoolest laialdaselt kasutatav. Nimelt saadavad modemid iga sideseansi alguses vastuvõtjas täpselt teada oleva treeningsignaali. Seda treeningsignaali kasutab vastuvõtja oma ekvalaiser koefitsientide adapteerimiseks. Peale treeningperioodi lõppu lülitub algoritm ümber selliselt, et soovitud signaalina kasutatakse vastuvõetud sümbolitest genereeritud signaali, mis on müravaba ja suure tõenäosusega võrdne saatja väljundsignaaliga. Seda selleks, et jälgida kanali impulsskarakteristiku võimalikke aeglasi muutusi. Sellist ekvalaiserit kutsutakse otsuseid tagasisidestavaks ekvalaiseriks (Decision Feedback Equalizer).

Normaliseeritud LMS algoritm

Eelnevalt nägime, et LMS algoritmi sammu suurus on otsustava tähtsusega algoritmi koonduvuse garanteerimiseks. Samuti leidsime, et koondumise garanteerimiseks on vajalik sisendsignaali korrelatsioonimaatriksi omaväärtuste või vähemalt sisendsignaali võimsuse teadmine. Mõnede

rakenduste korral võib selline informatsioon seadet konstrueerival inseneril olla. Samal ajal on rakendusi, kus sellise informatsiooni omamine ei ole võimalik. Näiteks juhul, kui adaptiivse filtri sisendsignaali on kõnesignaali, mille autokorrelatsioon ja võimsus muutuvad ajas kiiresti. Sisendsignaali võimsus ei ole seadme konstrueerimise ajal teada ka juhtudel, kus sisendsignaali läbib enne adaptiivse algoritmi jõudmist tundmatu süsteemi. Selles peatükis vaatleme normaliseeritud LMS algoritmi, mille sammu suurus muutub vastavalt sisendsignaali nii, et kindlustada adekvaatne töö erinevates olukordades.

Vastavalt valemile (5.30) koondub LMS algoritm, kui sammu suurus on valitud vahemikus

$$0 < \mu < \frac{2}{\sum_{k=0}^{M-1} E[x(n-k)]^2} \quad (5.37)$$

Normaliseeritud LMS algoritm jälgib LMS algoritmi üldist loogikat, kasutades selle võrratuse parema poole hetkehinnangut, seega sammu, mille pikkus on normaliseeritud sisendsignaali vektori eukleidilise normi ehk sisendsignaali vektori elementide ruutude summaga

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \frac{\alpha}{\|\mathbf{x}(n)\|^2} \mathbf{x}(n)e^*(n) \quad (5.38)$$

Seega võime vaadelda normaliseeritud LMS algoritmi, kui üht ajas muutuva sammuga LMS algoritmi varianti. Saab näidata, et normaliseeritud LMS algoritmi koondumiseks peab kehtima

$$0 < \alpha < 2 \quad (5.39)$$

ning kiireim koondumus saavutatakse, kui $\alpha = 1$. Praktikas valitakse α siiski tavaliselt ühest väiksemana, kuna nii nagu LMS algoritmi puhul, nii ka normaliseeritud LMS algoritmi korral kaasneb suure sammuga suur adaptatsiooni viga väljakujunenud režiimis.

Alternatiivselt on normaliseeritud LMS algoritmi võimalik tuletada ka järgmise piiranguga optimeerimisülesande lahendina [5].

$$\min \|\hat{\mathbf{w}}(n+1) - \hat{\mathbf{w}}(n)\|^2, \quad \text{piiranguga} \quad \hat{\mathbf{w}}^H(n+1)\mathbf{x}(n) = d(n) \quad (5.40)$$

Seega muudab normaliseeritud LMS igal sammul kaalude vektorit, võtmaks arvesse uusi andmeid nii vähe kui võimalik (eukleidilise normi mõttes).

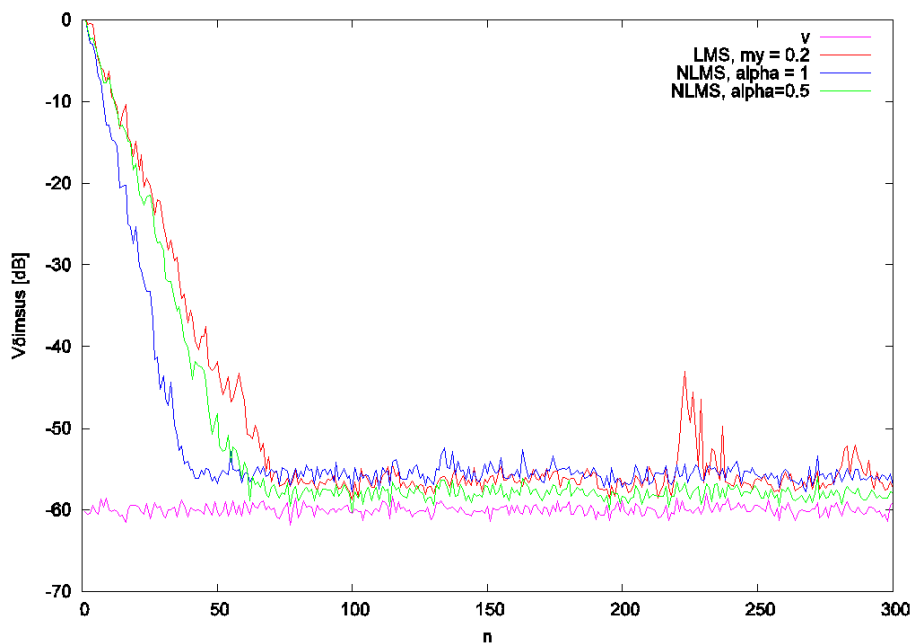
Praktikas on normaliseeritud LMS algoritmile mõistlik lisada nimetajasse konstantne liige β . Kui mitte muud, siis aitab see liige vähemalt vältida nulliga jagamist. Mitmed autorid soovivad selle liikme valida proportsionaalsena mõõtemärgi võimsusega vähendamaks müra mõju algoritmi tööle juhul, kui sisendsignaali on väike. Normaliseeritud LMS algoritmi lõplikuks kujuks saame siis

$$\begin{aligned}
y(n) &= \hat{\mathbf{w}}^H(n)\mathbf{x}(n) \\
e(n) &= d(n) - y(n) \\
\hat{\mathbf{w}}(n+1) &= \hat{\mathbf{w}}(n) + \frac{\alpha}{\beta + \|\mathbf{x}(n)\|^2} \mathbf{x}(n)e^*(n)
\end{aligned}
\tag{5.41}$$

Eeltoodud rekursioonis oleme tähistanud sisendsignaali vektori eukleidilise normi ruutu kui $\|\mathbf{x}(n)\|^2 = \mathbf{x}^H(n)\mathbf{x}(n)$.

Näide

Vaatleme taas süsteemi identifitseerimise näidet. Joonis 28 kujutab antud ülesannet skemaatiliselt. Nagu ennegi, on identifitseeritav süsteem lõpliku impulsskarakteristikuga filter ülekandefunktsiooniga $1 - az^{-1} + a^2z^{-2}$, mille identifitseerimiseks kasutame adaptiivset filtrit koefitsientide vektoriga \mathbf{w} . Olgu $a = 0,2$. Sisendsignaali $x(n)$ on valge müra võimsusega 1. Identifitseeritava süsteemi väljundile lisandub valge müra $v(n)$ võimsusega 10^{-6} . Protsessid x ja v on statistiliselt sõltumatud nullise keskvärtusega juhuslikud protsessid, mis alluvad normaaljaotusele. Kasutame LMS algoritmi sammu suurusega $\mu = 0,2$ ning kahte NLMS algoritmi vastavalt sammu suurustega $\alpha = 1$ ja $\alpha = 0,5$. Olgu meie adaptiivses filtris w , $M = 4$ koefitsienti algväärtustega $\mathbf{w}(0) = \mathbf{0}$.



Joonis 43. Õppekõverad

Joonis 43 kujutab näites kasutatud algoritmide õppekõveraid. Me näeme, et LMS algoritm koondub pisut aeglasemalt kui NLMS algoritmid. Samas on LMS algoritmi viga väljakujunenud režiimis võrreldav veaga, mille saavutab NLMS algoritm sammuga 1. Seda, et LMS algoritmi sammu edasine suurendamine ei ole mõistlik, võib näha ka väljavisetest, mis väljakujunenud režiimis esinevad. Seega võime järeldada, et normaliseeritud LMSil on (vähemalt teatud juhtudel) normaliseerimata variandiga võrreldes eeliseid ka statsionaarse sisendprotsessi korral.

LMS algorimi variandid

LMS algoritm on praktikas laialt levinud ning see on tinginud peale normaliseeritud LMS-i veel mitmete variantide tekke. Järgnevalt vaatlemegi mõnda neist lühidalt.

Esimeseks selliseks variatsiooniks oleks märgi LMS, kus veasignaali, sisendsignaali või mõlemad korraga asendatakse LMS kaalude arvutamise valemis vastava signaali märgiga nii, et vastav signaal võtab ainult kahte võimalikku väärtust +1 või -1. Matemaatiliselt võib nende algoritmide kaalude rekursioonid esitada kui

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \mu \mathbf{x}(n) \text{sign}(e(n)) \quad (5.42)$$

või

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \mu \text{sign}(\mathbf{x}(n))e(n) \quad (5.43)$$

või

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \mu \text{sign}(\mathbf{x}(n))e(n) \quad (5.44)$$

Kus $\text{sign}(\cdot)$ tähistab märgifunktsiooni, mille väärtuseks on +1, kui tema argument on positiivne ja -1, kui tema argument on negatiivne ning 0, kui tema argument on nulline. Komplexse argumenti korral tuleb võtta märgifunktsioon argumenti reaali- ja imaginaariosadest eraldi. Sellist LMS algoritmi modifikatsiooni kasutatakse vahel püsikomaga aritmeetikaga protsessorites, selleks et teha algoritmi robustsemaks suurte müra väljavisete vastu, mis muidu võiksid lahendit olulisel määral kallutada. Asja olemus seisneb selles, et kuna püsikomaga aritmeetika puhul on kõik signaalid esitatud täisarvudena, annab märgifunktsioon vähima võimaliku nihke soovitud suunas. On ütlematagi selge, et selliste algoritmide koondumiskiirus jääb standardse LMS algoritmi koondumiskiirusele oluliselt alla, kuid ta on suuteline toimima vägagi agressiivsete mürade korral.

Järgmise variandina vaatleksime algoritmi, mis saadakse analoogiliselt LMS-le, minimeerides vea ruudu asemel vea neljandat astet. Selliselt saadud algoritmi võib kutsuda vähima vea keskmise neljanda astme algoritmiks (inglise keeles Least Mean Fourth). Nagu LMS algoritmi puhulgi kasutame algoritmi tähistamiseks ingliskeelset lühendit LMF. Nii toimides saame kaalude rekursiooniks

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \mu \mathbf{x}(n) e^3(n) \quad (5.45)$$

Ambitsiooniks on saavutada paremad koonduvusomadused ning väiksem viga kui standardset LMS algoritmi kasutades. Tuleb aga öelda, et, vastupidiselt märgi LMS-le muudab vea neljanda astme minimeerimine algoritmi igasugustele müradele väga tundlikuks. Algoritmi võib vaadelda kui muutuva sammu suurusega LMS algoritmi, kus sammu suurus avaldub valemiga

$$\mu_{LMF} = \mu e^2(n) \quad (5.46)$$

Kirjandusest võib leida ka segu LMF-st ning standardsest LMS-st, mis annab nn segatud normiga LMS algoritmi

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \mu \mathbf{x}(n) (\delta e(n) + (1-\delta) e^3(n)) \quad (5.47)$$

Järgmiseks vaatleme nn lekkivat LMS algoritmi, mille saame, minimeerides kriteeriumifunktsiooni

$$J(n) = |e(n)|^2 + \alpha \|\mathbf{w}(n)\|^2 \quad (5.48)$$

kus $\|\mathbf{w}(n)\|^2 = \mathbf{w}^H \mathbf{w}$ on kaalude vektori eukleidilise normi ruut.

Nagu valemist (5.48) järeldub, soovitakse antud juhul saada vähimat veasignaali ruutkeskmist viga ning ka väikest kaalude vektori normi, mis tähendab väikseid filtrikoefitsiente. Kuivõrd olulised need kaks soovi suhtes teineteisega on, väljendab positiivne koefitsient α . Tavaliselt soovitakse siiski saada pigem väikese ruutkeskmise veaga lahendit kui väikest lahendit ning seega valitakse enamasti väike α . Lahendades toodud minimeerimisülesande, saame kaalude rekursiooniks

$$\hat{\mathbf{w}}(n+1) = (1 - \mu\alpha)\hat{\mathbf{w}}(n) + \mu\alpha\mathbf{x}(n)e(n), \quad 0 < (1 - \mu\alpha) < 1 \quad (5.49)$$

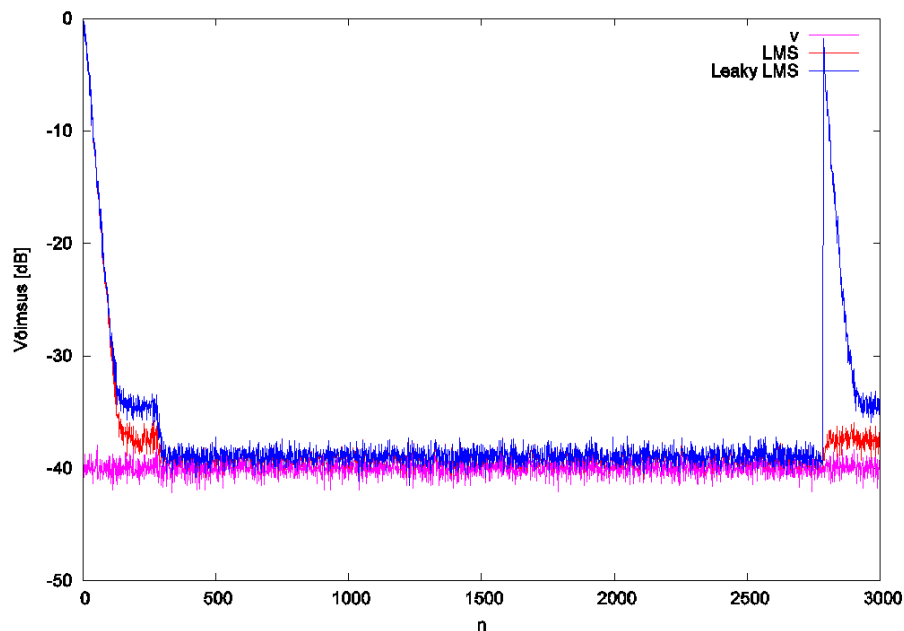
Toodud võrrand on väga sarnane vastava LMS võrrandiga (5.6). Vahe on ainult selles, et esimene liidetav on korrutatud teguriga $1 - \mu\alpha$. Tavaliselt on konstandid valitud nii, et $\mu\alpha$ on palju väiksem kui üks ja tegur $1 - \mu\alpha$, millega eelmise sammu filtri koefitsiente korrutatakse, on ainult pisut ühest väiksem. Selline korrutamisosoperatsioon viib lahendi aeglaselt nulli suunas. Siit ka algoritmi nimi – lekkiv LMS. Algoritmi selline konstruktsioon põhjustab nende moodide, mida sisendsignaal ei erguta, aeglase unustamise.

Erinevalt LMS-st ei ole lekkiva LMS algoritmi kaalude matemaatiline ootus väljakujunenud režiimis võrdne Wieneri filtriga. Selle asemel koondub algoritm nihkega Wieneri filtriks, kus hinnangu nihe on määratud konstandiga α .

$$\lim_{n \rightarrow \infty} E[\hat{\mathbf{w}}(n)] = (\mathbf{R} + \alpha\mathbf{I})^{-1} \mathbf{p} \quad (5.50)$$

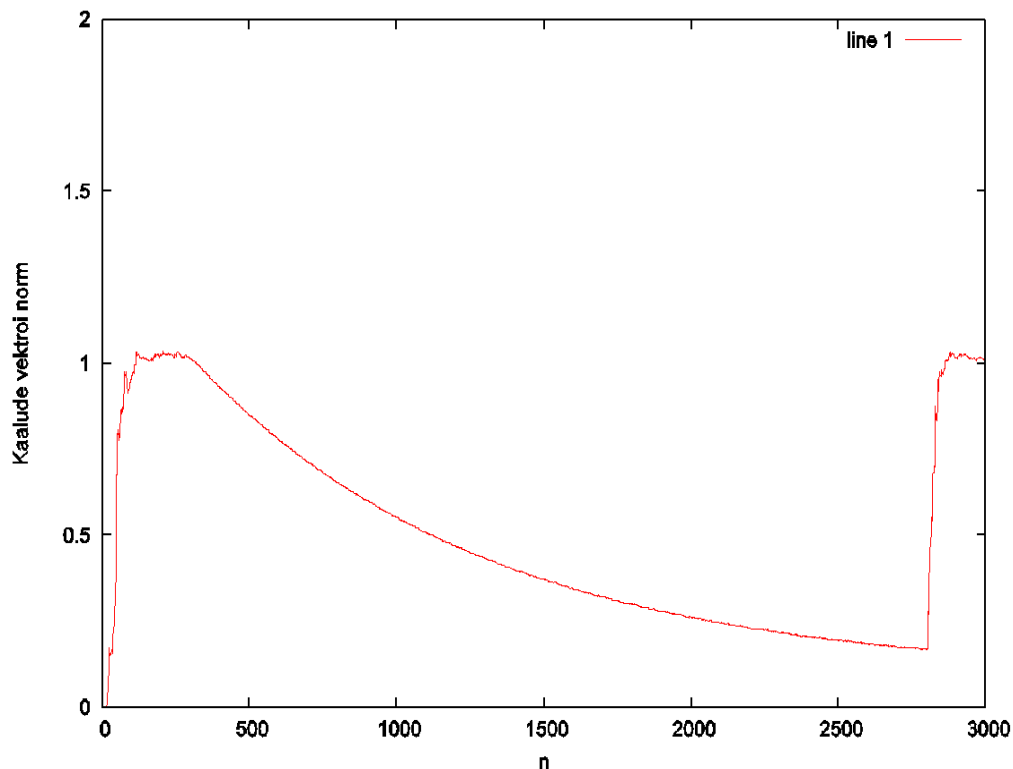
Näide

Vaatleme taas süsteemi identifitseerimise näidet (Joonis 28). Olgu seekord sisendsignaals ajaloigul $0 < n < 300$ valge müra võimsusega 1, vahemikus $301 < n < 2800$ siinussignaal $x(n) = \sin\left(2\pi\frac{n}{8}\right)$ ning alates diskreedist $n = 2801$ muutub sisendsignaal taas ühikulise võimsusega valgeks müraks.



Joonis 44. LMS ja lekkiva LMS algoritmi õppimiskõverad

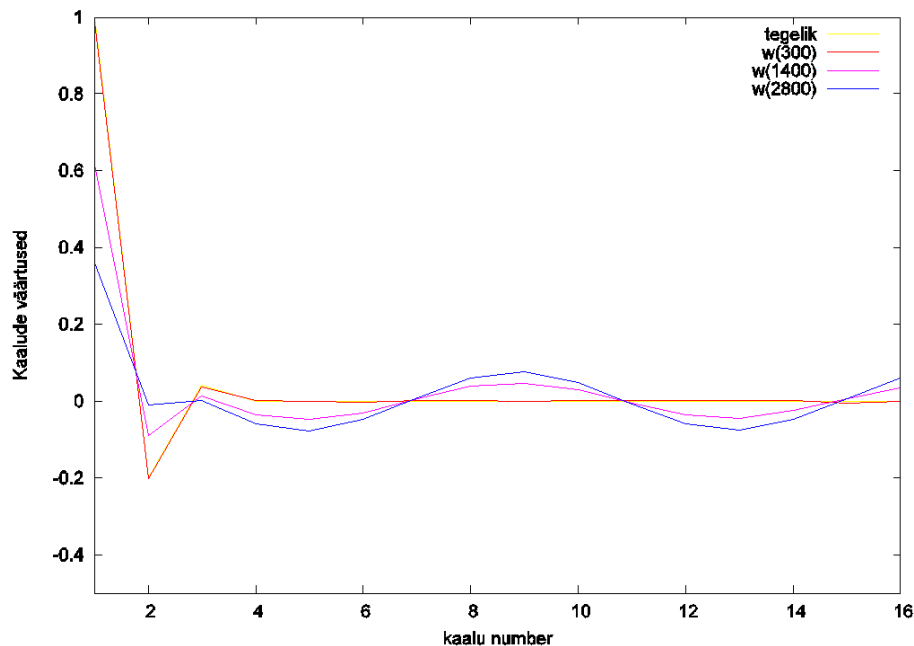
Joonis 44 kujutab LMS algoritmi ja lekkiva LMS algoritmi üle 100 realisatsiooni keskmistatud õppekõveraaid kirjeldatud sisendsignaali puhul. Sammu suuruseks on mõlema algoritmi korral $\mu = 0,05$ ja lekkiva LMS algoritmi lekkimise kiirust kontrolliv parameeter on valitud $\alpha = 0,01$, mis annab $1 - \mu\alpha = 0,0005$. Mõõtemüraks $v(n)$ on valge müra võimsusega 10^{-4} ja identifitseerimiseks on kasutatud 16 kaaluga adaptiivset filtrit. Nagu jooniselt näha, koonduvad mõlemad algoritmid esimese 300 iteratsiooni jooksul, kusjuures lekkiva LMS algoritmi viga väljakujunenud režiimis jääb veidi suuremaks standardse LMS-i veast. Ajal, mil sisendsignaaliks on sinusoid, on algoritmide vea võimsus ligikaudu võrdne. Kuid kui sisendsignaal muutub uuesti valgeks müra diskreedil 2801, on lekkiv LMS unustanud palju tundmatu süsteemi kohta esialgselt õpitust ja algoritm peab uuesti koonduma.



Joonis 45. Kaalude vektori normi areng ajas

Joonis 45 näitab, mis juhtub katse jooksul kaalude vektori normiga. Ajavahemikus $0 < n < 300$ kaalude vektori norm kasvab filtri koondudes, stabiliseerudes tasemele $\sqrt{1 + \alpha^2 + \alpha^4} = 1,02$ lähedal, seega taseme lähedal, mis on määratud identifitseeritava süsteemi poolt. Järgnevas ajahemikus $300 < n < 2800$ hakkab kaalude vektori norm aeglaselt vähenema, kuna selles ajavahemikus on sisendsignaaliks üksainus sinusoid. Üheainsa sinusoidi abil on aga võimalik määrata identifitseeritava süsteemi sageduskarakteristik ühesainsas punktis – sagedusel, mis langeb kokku sinusoidi sagedusega. Vaadeldes sisendsignaali korrelatsioonimaatriksit, näeme, et tal on ainult üks nullist erinev omaväärtus. Selline sisendsignaal ei ole süsteemi täielikuks identifitseerimiseks küllaldaselt ergutav ning Wieneri filtrit määrav võrrandisüsteem on singulaarne. Sellisel juhul on normaalvõrranditel lõpmata palju lahendeid ning adaptiivne filter võib valida neist ükskõik millise. Lekkiva adaptiivse algoritmi korral hakkabki lahendus kalduma vähima võimaliku normiga lahenduse suunas, unustades esialgselt õpitud süsteemi sageduskarakteristiku kõigil sagedustel peale sinusoidi sageduse. Samal ajal, nagu nähtub eelmiselt jooniselt, on veasignaali võimsus ligikaudu konstantne. Ajavahemikus $n >$

2800 on sisendsignaalsiks taas valge müra ning meie adaptiivne algoritm koondub uuesti lahenduseks, mis katab kogu sagedusala.



Joonis 46. Kaalude areng

Joonis 46 kujutab adaptiivse filtri kaale adaptatsiooni erinevatel ajahetkedel. Näeme, et ajahetkeks $n = 300$ on adaptiivse filtri kaalud jõudnud korrektse lahendi lähetele ja ei ole toodud joonisel korrektsest lahendist eristatavad. Järgnevatel ajahetkedel $n = 1400$ ja $n = 2800$ kaalud teisenevad. Esimene, kõige suurem koefitsient väheneb ning kogu kaalude vektor muutub järjest sarnasemaks sisendsignaali võrdse sagedusega sinusoidiga.

Plokk LMS on LMS algoritmi variant, kus kaalude vektorit ei uuendata pärast iga uue sisend- ja soovitud signaali diskreedi saabumist, vaid ainult üks kord iga L diskreedi korral

$$\hat{\mathbf{w}}(k+1) = \hat{\mathbf{w}}(k) + \mu \sum_{i=0}^{L-1} \mathbf{x}(kM+i) e^*(kM+i) \quad (5.51)$$

kasutades vastavalt üle L diskreedi keskmistatud gradienti.

$$\hat{\nabla}(k) = -\frac{2}{L} \sum_{i=0}^{L-1} \mathbf{x}(kM+i) e^*(kM+i) \quad (5.52)$$

Algoritm koondub kui

$$0 < \mu < \frac{2}{L \text{Tr}(\mathbf{R})} \quad (5.53)$$

Seega tuleb algoritmi samm valida L korda vastava standardse LMS sammust väiksem. Tuleb aga meeles pidada, et kuna kasutusel on keskmistatud gradient, liigub lahend Wieneri filtri poole mööda sirgemat teed kui standardse LMS algoritmi lahend, mis kompenseerib osaliselt väiksemast lubatud sammust tingitud aeglasemat koondumiskiirust. ¶

Eelnevalt töime ära ühe võimaliku LMS algoritmi analüüsi. Nagu juba märgitud, ei ole toodud analüüs kaugeltki ainus võimalik omataoline. Nimelt, tuginedes nn energia säilitamise printsiibile, on võimalik algoritmide perekonda

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \mu \mathbf{x}(n) f(e(n)) \quad (5.54)$$

kus $f(\cdot)$ tähistab mingit funktsiooni, edukalt analüüsida, saades järgnevad tulemused [19] (toome nad siinkohal ära ilma tuletusteta).

	$f(e(n))$	$J_{ex}(\infty)$
LMS	$e(i)$	$\frac{\mu \sigma_v^2 \text{Tr}(\mathbf{R})}{2 - \mu \text{Tr}(\mathbf{R})}$
NLMS	$\frac{e(i)}{\ \mathbf{x}(i)\ ^2}$	$\frac{\mu \sigma_v^2}{2 - \mu}$
Märgi LMS	$\text{sign}[e(i)]$	$\frac{\alpha}{2} \left(\alpha + \sqrt{\alpha^2 + 4\sigma_v^2} \right), \quad \alpha = \sqrt{\frac{\pi}{8}} \mu \text{Tr}(\mathbf{R})$
LMF	$e^3(i)$	$\frac{\mu \xi_6^v \text{Tr}(\mathbf{R})}{4\sigma_v^2 - 9\mu \xi_4^v \text{Tr}(\mathbf{R})}, \quad \xi_4^v = E(v(i) ^4), \quad \xi_6^v = E(v(i) ^6)$

6. Vähimruutude meetod

Vähimruutude meetod on laialdaselt rakendatav meetod sobitamaks katsetulemusi mingit nähtust kirjeldava matemaatilise mudeliga. Mudeli headuse kriteeriumiks on vigade ruutude summa, mida käesolev meetod minimeerib. Sellest ka nimetus – vähimruutude meetod. Nii nagu Wieneri filtri puhulgi, ei ole vähimruutude meetodil saadav lahend iseenesest adaptiivne, kuid ta on vajalik valmistamiseks ette mõnede adaptiivsete algoritmide, näiteks rekursiivse vähimruutude meetodi käsitlust järgmistes peatükkides. Wieneri teooria ja vähimruutude meetodi erinevuseks on järgmine asjaolu. Wieneri teooria kasutab üle ansambli keskmistatud korrelatsioonimaatrikseid ja vektoreid ning seega on tema abil saadud lahendid optimaalsed tõenäosuslikus tähenduses kõigi protsessi realisatsioonide korral. Vähimruutude meetod keskendub aga konkreetsele protsessi realisatsioonile ning kasutab keskmistamist ajas, leides nii probleemile deterministliku lahendi.

Olgu meil füüsikaline nähtus, mis on kirjeldatav kahe andmehulga d ja x abil. $d(i)$ on muutuja, mida katse käigus otseselt mõõdetakse ning ta on eelduse kohaselt reaktsioon sisenditele $x(i), \dots, x(i-M+1)$. Hüpoteesi kohaselt on d funktsionaalne sõltuvus sisenditest $x(i), \dots, x(i-M+1)$ lineaarne. Seega avaldub meie mõõdetav suurus sisendite kaalutud summana

$$d(i) = \sum_{k=0}^{M-1} w_{o,k}^* x(i-k) + e_o(i) \quad (6.1)$$

Kaalud w on mudeli tundmatud parameetrid ja $e_o(i)$ on mõõteviga. Alumine indeks o väljendab optimaalsust vähimruutude mõttes. Selline mudel on tuntud lineaarse regressiooni nime all. Eeldame, et mõõteviga $e_o(i)$ on valge, nullise keskväärtusega juhuslik protsess. Vähimruutude meetod leiab regressioonikoefitsiendid w , minimeerides mõõtevigade ruutude summat

$$V = \sum_{i=i_1}^{i_2} |e(i)|^2 = \sum_{i=i_1}^{i_2} \left| d(i) - \sum_{k=0}^{M-1} w_{ok}^* x(i-k) \right|^2 \quad (6.2)$$

Võrreldes eeltoodud Wieneri teooria kriteeriumifunktsiooniga (3.4), näeme, et vähimruutude meetodi kriteeriumifunktsioon ei sisalda matemaatilisi ootusi, nagu seda teeb Wieneri teooria. Selle asemel on kriteeriumiks lõpliku hulga vigade ruutude summa.

Läheme nüüd oma tähistustes üle maatrikskujule. Selleks koondame oma sisendandmed x maatriksisse \mathbf{X} , millel on mudeli pikkusega M võrdne arv ridu ning vastavalt $N - M$ veergu.

$$\mathbf{X} = \begin{bmatrix} x(M) & x(M+1) & \dots & x(N) \\ x(M-1) & x(M) & \dots & x(N-1) \\ \vdots & \vdots & \ddots & \vdots \\ x(1) & x(2) & \dots & x(N-M+1) \end{bmatrix} \quad (6.3)$$

Vähimruutude meetod vaatleb maailma nagu läbi teatud suurusega akna. Valemis (6.2) on erinevad aknad määratud summeerimisradade i_1 ja i_2 valikuga. Valides $i_1=M$ ja $i_2=N$, saame nn kovariatsioonimeetodi, mille andmete maatriks \mathbf{X} ongi esitatud valemiga (6.3). Selline valik eeldab, et me ei tea andmetest väljaspool akent mitte midagi. Selline lähenemine on praktikas kõige tavalisem, kuid mitte ainus võimalik. Näiteks on võimalik ka eeldada, et kõik andmed väljaspool meie akent on võrdsed nulliga, ning täiendades vastavalt maatriksit \mathbf{X} nende nullidega saame nn korrelatsioonimeetodi.

Koondades mõõtetulemused veeruvektorisse \mathbf{d} , mudeli optimaalsed kaalud teise veeruvektorisse \mathbf{w}_o , ning optimaalse mudeli jääkvead kolmandasse veeruvektorisse \mathbf{e}_o , võime võrrandi (6.1) esitada kompaktselt kujul

$$\mathbf{d} = \mathbf{X}^H \mathbf{w}_o + \mathbf{e}_o \quad (6.4)$$

Kasutades äsjaleitud võrrandit (6.4), võime viia maatrikskujule ka võrrandi (6.2), saades kriteeriumifunktsiooni jaoks avaldise

$$\begin{aligned} V &= \sum_{i=i_1}^{i_2} |e(i)|^2 = \sum_{i=i_1}^{i_2} \left| d(i) - \sum_{k=0}^{M-1} w_k^* x(i-k) \right|^2 = \mathbf{e}^H \mathbf{e} \\ &= (\mathbf{d} - \mathbf{X}^H \mathbf{w})^H (\mathbf{d} - \mathbf{X}^H \mathbf{w}) = \mathbf{d}^H \mathbf{d} - \mathbf{d}^H \mathbf{X}^H \mathbf{w} - \mathbf{w}^H \mathbf{X} \mathbf{d} + \mathbf{w}^H \mathbf{X} \mathbf{X}^H \mathbf{w} \end{aligned} \quad (6.5)$$

Optimaalsete kaalude leidmiseks võtame eeltoodud kriteeriumist tuletise kaalude vektori \mathbf{w} järgi ja võrdsustame ta nulliga

$$\frac{\partial}{\partial \mathbf{w}^*} V = 0 \Rightarrow -\mathbf{X} \mathbf{d} + \mathbf{X} \mathbf{X}^H \hat{\mathbf{w}} = 0 \quad (6.6)$$

Seega on vähimruutude meetodi normaalvõrranditeks

$$\mathbf{XX}^H \hat{\mathbf{w}} = \mathbf{Xd} \quad (6.7)$$

Eeldusel, et maatriks \mathbf{XX}^H ei ole singulaarne, st maatriksi \mathbf{XX}^H pöördmaatriks eksisteerib, on võrrandi (6.7) lahendiks

$$\hat{\mathbf{w}} = (\mathbf{XX}^H)^{-1} \mathbf{Xd} \quad (6.8)$$

Asendades saadud lahendi vea ruutude summa avaldise (6.5), saame vähimruutude meetodi vea jaoks avaldise

$$\begin{aligned} V_{\min} &= \mathbf{d}^H \mathbf{d} - \mathbf{d}^H \mathbf{X}^H \hat{\mathbf{w}} - \hat{\mathbf{w}}^H \mathbf{Xd} + \hat{\mathbf{w}}^H \mathbf{XX}^H \hat{\mathbf{w}} \\ &= \mathbf{d}^H \mathbf{d} - \mathbf{d}^H \mathbf{X}^H (\mathbf{XX}^H)^{-1} \mathbf{Xd} - \mathbf{d}^H \mathbf{X}^H (\mathbf{XX}^H)^{-1} \mathbf{Xd} + \mathbf{d}^H \mathbf{X}^H (\mathbf{XX}^H)^{-1} \mathbf{XX}^H (\mathbf{XX}^H)^{-1} \mathbf{Xd} \\ &= \mathbf{d}^H \mathbf{d} - \mathbf{d}^H \mathbf{X}^H (\mathbf{XX}^H)^{-1} \mathbf{Xd} \end{aligned} \quad (6.9)$$

Näide

Olgu $d(1), \dots, d(N)$ mingi konstantse suuruse c korduval mõõtmisel saadud tulemused ning olgu meie ülesandeks leida suuruse c vähimruutude mõttes parim hinnang. Valime valemis (6.4) esinevad vektorid ja maatriksid järgmiselt:

$$\mathbf{w} = c$$

$$\mathbf{d} = [d_1 \quad d_2 \quad \dots \quad d_N]^T \quad (6.10)$$

$$\mathbf{X} = [1 \quad 1 \quad \dots \quad 1]$$

Kuna otsime ainult ühte kaalu, on meie mudeli suurus $M = 1$ ja seega on otsitav kaalude vektor ühedimensionaalne muutuja c . Kõik N mõõtetulemust on koondatud vektorisse \mathbf{d} ning sisendi maatriks \mathbf{X} koosneb ainult ühtedest (kuna tegelikult ülesandes sisendit ei ole ja me mõõdame otseselt uuritavat suurust ennast, mitte tema sõltuvust millestki). Vähimruutude normaalvõrrandid (6.7) võime siis antud ülesande jaoks kirja panna kujul

$$[1 \quad 1 \quad \dots \quad 1] \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} c = [1 \quad 1 \quad \dots \quad 1] \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{bmatrix} \quad (6.11)$$

Millest tuleneb, et

$$Nc = \sum_{i=1}^N d_i \Rightarrow c = \frac{1}{N} \sum_{i=1}^N d_i \quad (6.12)$$

Seega jõudsimme igati loogilise tulemuseni – vähimruutude hinnang korduvalt mõõdetud suurusele on mõõtetulemuste aritmeetiline keskmine.

Vähimruutude meetodi normaalkvõrrandid (6.7) on oma kujult sarnased Wieneri teooria normaalkvõrranditega (3.26). Rõhutamaks seda sarnasust veelgi võime defineerida ajas keskmistatud korrelatsioonimaatriksi

$$\Phi = \mathbf{X}\mathbf{X}^H = \sum_{i=M}^N \mathbf{x}(i)\mathbf{x}^H(i), \quad (6.13)$$

kus $\mathbf{x}(i) = [x(i) \ x(i-1) \ \dots \ x(i-M)]^T$ on maatriksi \mathbf{X} i-s veerg. Sarnaselt võime defineerida ka ajas keskmistatud ristkorrelatsiooni vektori

$$\mathbf{z} = \mathbf{X}\mathbf{d} = \sum_{i=M}^N \mathbf{x}(i)d^*(i) \quad (6.14)$$

Vähimruutude meetodi normaalkvõrrandid võib nüüd kirjutada kujul

$$\Phi \hat{\mathbf{w}} = \mathbf{z} \quad (6.15)$$

Ajas keskmistatud korrelatsioonimaatriksi omadused

Omadus 1

Ajas keskmistatud korrelatsioonimaatriksil on Hermiti sümmeetria

$$\Phi^H = \Phi \quad (6.16)$$

See omadus tuleneb otseselt ajas keskmistatud korrelatsioonimaatriksi definitsioonist (6.13).

Omadus 2

Ajas keskmistatud korrelatsioonimaatriks on mittenegatiivselt määratud, st mis tahes vektori \mathbf{u} korral kehtib

$$\mathbf{u}^H \Phi \mathbf{u} \geq 0. \quad (6.17)$$

Tõepoolest, kasutades ajas keskmistatud korrelatsioonimaatriksi määratlust, saame

$$\begin{aligned} \mathbf{u}^H \Phi \mathbf{u} &= \sum_{i=M}^N \mathbf{u}^H \mathbf{x}(i)\mathbf{x}^H(i)\mathbf{u} \\ &= \sum_{i=M}^N \mathbf{u}^H \mathbf{x}(i) [\mathbf{u}^H \mathbf{x}(i)]^H \\ &= \sum_{i=M}^N |\mathbf{u}^H \mathbf{x}(i)|^2 \geq 0 \end{aligned} \quad (6.18)$$

Omadus 3

Tulenevalt mittenegatiivsest määratusest on kõik ajas keskmistatud korrelatsioonimaatriksi omaväärtused reaalsed ja mittenegatiivsed.

Omadus 4

Ajas keskmistatud korrelatsioonimaatriks ei ole Toeplitzi maatriks, vaid avaldub kahe Toeplitzi maatriksi, mis on teineteise Hermiti transponaadid, \mathbf{X}^H ja \mathbf{X} korrutisena. Üldjuhul ei ole ajas

keskmistatud korrelatsioonimaatriks Toeplitzi maatriks. See on oluline erinevus Wieneri teooriast, kus korrelatsioonimaatriks on definitsiooni põhjal alati Toeplitzi maatriks.

Projektsiooni operaator

Vaatleme nüüd vähimruutude hinnangut jälgitud andmete vektorile \mathbf{d} . Selleks hinnanguks on sisendandmete maatriksi \mathbf{X}^H korrutis vähimruutude meetodil leitud regressioonikoefitsientide vektoriga $\hat{\mathbf{w}}$,

$$\hat{\mathbf{d}} = \mathbf{X}^H \hat{\mathbf{w}} = \mathbf{X}^H (\mathbf{X}\mathbf{X}^H)^{-1} \mathbf{X}\mathbf{d} \quad (6.19)$$

Teine võrdus eeltoodud valemis saadakse, kui asendada vähimruutude kaalud normaalvõrrandi lahendi avaldisest (6.8). Viime nüüd sisse tähistuse

$$\mathbf{P} = \mathbf{X}^H (\mathbf{X}\mathbf{X}^H)^{-1} \mathbf{X} \quad (6.20)$$

Maatriksit \mathbf{P} kutsutakse projektsioonimaatriksiks. Nimetus tuleneb sellest, et projektsioonimaatriksiga vasakult poolt korrutamine projekteerib vektori \mathbf{d} maatriksi \mathbf{X} veergude poolt määratud alamruumi, andes tulemuseks vastavalt valemile (6.19) andmete vektori hinnangu. Maatriks

$$\mathbf{I} - \mathbf{P} = \mathbf{I} - \mathbf{X}^H (\mathbf{X}\mathbf{X}^H)^{-1} \mathbf{X} \quad (6.21)$$

on maatriksi \mathbf{P} ortogonaalne täiend ning tema rakendamine vektorile \mathbf{d} annab hinnangu vea.

$$\mathbf{e}_{\min} = (\mathbf{I} - \mathbf{P})\mathbf{d} = \mathbf{d} - \mathbf{X}^H (\mathbf{X}\mathbf{X}^H)^{-1} \mathbf{X}\mathbf{d} = \mathbf{d} - \hat{\mathbf{d}} \quad (6.22)$$

Nii nagu Wieneri filtrite korral, kehtib ka vähimruutude meetodi puhul ortogonaalsusprintsip. Erinevalt Wieneri filtrite teooriast, kus ortogonaalsus defineeriti vektorite skalaarkorrutise matemaatilise ootuse kaudu, on vähimruutude meetodi kontekstis ortogonaalsus määratud vektorite skalaarkorrutisega. Seega ütleme, et kaks vektorit on ortogonaalsed, kui nende skalaarkorrutis on võrdne nulliga. Leides vähimruutude meetodil leitud andmete hinnangu vektori (6.19) ja hinnangu vea vektori (6.22) skalaarkorrutise, on kerge veenduda, et need vektorid on ortogonaalsed

$$\hat{\mathbf{d}}^H \mathbf{e}_{\min} = \mathbf{d}^H \mathbf{X}^H (\mathbf{X}\mathbf{X}^H)^{-1} \mathbf{X}\mathbf{d} - \mathbf{d}^H \mathbf{X}^H (\mathbf{X}\mathbf{X}^H)^{-1} \mathbf{X}\mathbf{X}^H (\mathbf{X}\mathbf{X}^H)^{-1} \mathbf{X}\mathbf{d} = 0 \quad (6.23)$$

Sellest tulenevalt võime, nagu Wieneri filtritegi puhul, öelda, et mõõdetud andmete hinnang on mõõdetud andmete projektsioon sisendvektorite poolt määratud alamruumi ning vea vektor on selle alamruumiga ortogonaalne signaali d komponent, mida ei ole võimalik sisendandmete abil seletada.

Vähimruutude hinnangu omadused

Omadus 1

Vähimruutude hinnang on nihketa hinnang eeldusel, et hinnangu viga $e_o(i)$ on nullise keskvaertusega juhuslik protsess.

Olgu meil vähimruutude meetodil leitud regressioonikoefitsientide vektori hinnang vastavalt valemile (6.8). Asendades sellesse hinnangusse jälgitud andmete vektori \mathbf{d} avaldise mudelist (6.4), saame

$$\hat{\mathbf{w}} = (\mathbf{X}\mathbf{X}^H)^{-1} \mathbf{X}\mathbf{d} = (\mathbf{X}\mathbf{X}^H)^{-1} \mathbf{X}\mathbf{X}^H \mathbf{w}_o + (\mathbf{X}\mathbf{X}^H)^{-1} \mathbf{X}\mathbf{e}_o = \mathbf{w}_o + (\mathbf{X}\mathbf{X}^H)^{-1} \mathbf{X}\mathbf{e}_o \quad (6.24)$$

Kuna \mathbf{X} on vaadeldav ja seega täpselt teada ning \mathbf{e}_o on nullise keskvaertusega valge juhuslik protsess, saame:

$$E(\hat{\mathbf{w}}) = \mathbf{w}_o \quad (6.25)$$

Saime tulemuseks, et vähimruutude hinnangu keskvaertus on võrdne tegeliku koefitsientide vektoriga mudelis (6.4). Seega on vähimruutude hinnang nihketa.

Omadus 2

Kui hinnangu viga $e_o(i)$ on nullise keskvaertusega valge juhuslik protsess dispersiooniga σ^2 , siis on vähimruutude hinnangu $\hat{\mathbf{w}}$ kovariatsioonimaatriks $\sigma^2 \mathbf{\Phi}^{-1}$.

Hinnangu vea kovariatsioonimaatriksi leidmiseks kasutame äsja saadud valemit (6.24) ning eeldust, et e_o on valge müra kovariatsioonimaatriksiga $\sigma^2 \mathbf{I}$, saades

$$\begin{aligned} E[(\hat{\mathbf{w}} - \mathbf{w}_o)(\hat{\mathbf{w}} - \mathbf{w}_o)^H] &= E[(\mathbf{X}^H \mathbf{X})^{-1} \mathbf{X}^H \mathbf{e}_o \mathbf{e}_o^H \mathbf{X} (\mathbf{X}^H \mathbf{X})^{-1}] \\ &= (\mathbf{X}^H \mathbf{X})^{-1} \mathbf{X}^H E[\mathbf{e}_o \mathbf{e}_o^H] \mathbf{X} (\mathbf{X}^H \mathbf{X})^{-1} = \sigma^2 (\mathbf{X}^H \mathbf{X})^{-1} = \sigma^2 \mathbf{\Phi}^{-1} \end{aligned} \quad (6.26)$$

Omadus 3

Vähimruutude meetodil leitud hinnang on parim võimalik lineaarne nihketa hinnang koefitsientide vektorile \mathbf{w}_o .

Olgu meil mingi suvaline nihketa \mathbf{w}_o hinnang, mis on saadud lineaarse kombinatsioonina andmete vektori \mathbf{d} elementidest. Tähistame selle suvalise nihketa hinnangu lainelise joonekesega vektori kohal. Mis tahes sellise lineaarse kombinatsiooni võime üles kirjutada andmete vektori \mathbf{d} ja mingi $M \times N-M+1$ maatriksi \mathbf{B} (mis määrab, millise lineaarse kombinatsiooniga konkreetselt on tegu) korrutisena

$$\tilde{\mathbf{w}} = \mathbf{B}\mathbf{d} \quad (6.27)$$

Asendades vektori \mathbf{d} mudelist (6.4), saame

$$\tilde{\mathbf{w}} = \mathbf{B}(\mathbf{X}^H \mathbf{w}_o + \mathbf{e}_o) = \mathbf{B}\mathbf{X}^H \mathbf{w}_o + \mathbf{B}\mathbf{e}_o \quad (6.28)$$

Kuna \mathbf{e}_o on eelduse põhjal nullise keskvaertusega, saame oma suvalise hinnangu matemaatiliseks ootuseks

$$E[\tilde{\mathbf{w}}] = \mathbf{B}\mathbf{X}^H \mathbf{w}_o \quad (6.29)$$

Eelnevast valemist selgub aga, et nihketa hinnangu saamiseks peab kehtima

$$\mathbf{B}\mathbf{X}^H = \mathbf{I} \quad (6.30)$$

Ning siis saame avaldisest (6.28)

$$\tilde{\mathbf{w}} = \mathbf{w}_o + \mathbf{B}\mathbf{e}_o \quad (6.31)$$

Leiame nüüd hinnangu vea kovariatsioonimaatriksi

$$E[(\tilde{\mathbf{w}} - \mathbf{w}_o)(\tilde{\mathbf{w}} - \mathbf{w}_o)^H] = E[\mathbf{B}\mathbf{e}_o \mathbf{e}_o^H \mathbf{B}^H] = \sigma^2 \mathbf{B}\mathbf{B}^H \quad (6.32)$$

Edasi vaatleme, millises vahekorras on äsja leitud suvalise nihketa hinnangu kovariatsioonimaatriks ning vähimruutude meetodil saadud hinnangu vea korrelatsioonimaatriks (6.26). Defineerime maatriksi

$$\Psi = \mathbf{B} - (\mathbf{X}\mathbf{X}^H)^{-1} \mathbf{X} \quad (6.33)$$

ja vaatleme tema korrutist iseenda Hermiti transponaadiga

$$\begin{aligned} \Psi\Psi^H &= \left[\mathbf{B} - (\mathbf{X}\mathbf{X}^H)^{-1} \mathbf{X} \right] \left[\mathbf{B}^H - \mathbf{X}^H (\mathbf{X}\mathbf{X}^H)^{-1} \right] \\ &= \mathbf{B}\mathbf{B}^H - \mathbf{B}\mathbf{X}^H (\mathbf{X}\mathbf{X}^H)^{-1} - (\mathbf{X}\mathbf{X}^H)^{-1} \mathbf{X}\mathbf{B}^H + (\mathbf{X}\mathbf{X}^H)^{-1} \mathbf{X}\mathbf{X}^H (\mathbf{X}\mathbf{X}^H)^{-1} \\ &= \mathbf{B}\mathbf{B}^H - (\mathbf{X}\mathbf{X}^H)^{-1} \end{aligned} \quad (6.34)$$

Selline maatriks on (analoogselt kompleksarvu mooduli ruuduga) alati mittenegatiivselt määratud. Mittenegatiivsest määratusest tulenevalt on maatriksi $\Psi\Psi^H$ kõik diagonaalelemendid mittenegatiivsed ning seega kehtib

$$\sigma^2 \text{diag}[\mathbf{B}\mathbf{B}^H] \geq \sigma^2 \text{diag}[(\mathbf{X}\mathbf{X}^H)^{-1}] \quad (6.35)$$

Valemi (6.32) põhjal on $\sigma^2 \mathbf{B}\mathbf{B}^H$ meie suvalise hinnangu kovariatsioonimaatriks ning vastavalt valemile (6.26) on $\sigma^2 (\mathbf{X}\mathbf{X}^H)^{-1}$ vähimruutude hinnangu kovariatsioonimaatriks. Seega on vähimruutude hinnang parim nihketa lineaarne hinnang, kuna kõik tema elemendid omavad väiksemat või võrdset dispersiooni kui mis tahes teise lineaarse hinnangu elemendid.

Väärrib märkimist, et kui e on valge protsess, mis allub normaaljaotusele, siis vähimruutude meetodi hinnang on parim mitte ainult lineaarsete nihketa hinnangute klassis, vaid üldiselt, ehk vea korrelatsioonimaatriks küünib Cramer-Rao tingimuste poolt määratud piirini

$$E[(\tilde{\mathbf{w}} - \mathbf{w}_o)(\tilde{\mathbf{w}} - \mathbf{w}_o)^H] \geq \mathbf{J}^{-1} \quad (6.36)$$

kus maatriks \mathbf{J} on Fisheri informatsioonimaatriks

$$\mathbf{J} = E \left[\left(\frac{\partial l}{\partial \mathbf{w}_o^*} \right) \left(\frac{\partial l}{\partial \mathbf{w}_o^T} \right) \right] \quad (6.37)$$

ning l on logaritmiline tõepärafunktsioon, mis on mõõtevigade vektori \mathbf{e}_o jaotustihedusfunktsiooni naturaallõgaritm

$$l = \ln f_E(\mathbf{e}_o) \quad (6.38)$$

Kui f_E on normaaljaotus, siis saab tõestada, et

$$\mathbf{J} = \frac{1}{\sigma^2} \mathbf{X}\mathbf{X}^H, \quad (6.39)$$

mis langeb kokku vähimruutude vea kovariatsioonimaatriksi (6.26) pöördmaatriksiga.

Lahutus singulaarväärtuste järgi

Vähimruutude meetodi hinnangu arvutamine otse valemist (6.8), mida siin lihtsuse huvides kordame,

$$\hat{\mathbf{w}} = (\mathbf{X}\mathbf{X}^H)^{-1} \mathbf{X}\mathbf{d} = \Phi^{-1}\mathbf{z} \quad (6.40)$$

eeldab korrelatsioonimaatriksi $\Phi = \mathbf{X}\mathbf{X}^H$ leidmist ja pööramist ning korrutamist eelnevalt leitud ristkorrelatsioonivektoriga $\mathbf{z} = \mathbf{X}\mathbf{d}$. See ei ole aga numbriliselt hea protseduur, kuna andmed tulevad rehendumisse sisse ruutudena ja see nõuab arvutuste läbiviimiseks kahekordset dünaamilist diapasooni ehk kaks korda rohkem bitte. Leides lahendi läbi singulaarväärtuste lahutuse õnnestub seda probleemi vältida. Defineerime maatriksi

$$\mathbf{X}^+ = (\mathbf{X}\mathbf{X}^H)^{-1} \mathbf{X} \quad (6.41)$$

siis

$$\hat{\mathbf{w}} = \mathbf{X}^+ \mathbf{d} \quad (6.42)$$

Maatriksit \mathbf{X}^+ kutsutakse maatriksi \mathbf{X}^H Moore-Penrose'i pseudopöördmaatriksiks. Praktikas sisaldab \mathbf{X}^H sageli üksteisest lineaarselt sõltuvaid veerge. Osutub, et pseudopöördmaatriks ja lahutus singulaarväärtuste järgi annavad meile võimsa matemaatilise tööriista, mis lubab lahendada nii ülemääratud kui ka alamääratud võrrandisüsteeme. See on vägagi oluline praktikas, kus võrrandisüsteem moodustub mõõdetud andmetest ning me ei tea, kui paljud maatriksi \mathbf{X} veerud on üksteisest lineaarselt sõltumatud.

Olgu meil võrrandisüsteem

$$\mathbf{X}^H \hat{\mathbf{w}} = \mathbf{d} \quad (6.43)$$

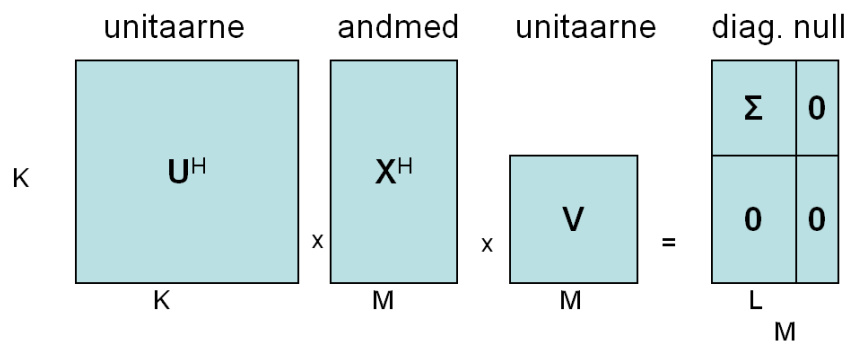
kus \mathbf{X}^H on $M \times K$ andmete maatriks, \mathbf{d} on $K \times 1$ vektor ja $\hat{\mathbf{w}}$ on $M \times 1$ hinnangu vektor. Korrutades vasakult \mathbf{X} -ga saaksime normaalvõrrandid. Osutub, et mis tahes \mathbf{X}^H puhul eksisteerivad kaks unitaarset maatriksit \mathbf{U} ja \mathbf{V} nii, et kehtib

$$\mathbf{U}^H \mathbf{X}^H \mathbf{V} = \begin{bmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (6.44)$$

kus Σ on diagonaalmaatriks, mille peadiagonaalil on maatriksi nn singulaarväärtused

$$\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_L) \quad (6.45)$$

ning need singulaarväärtused on järjestatud kahanevas järjekorras $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_L$. Joonis 47 illustreerib dekompositsiooni singulaarväärtuste järgi.



Joonis 47. Maatriksi dekompositsioon singulaarväärtuse järgi

Meie põhiline huvi singulaarväärtuste järgi dekompositsiooni vastu tuleneb soovist defineerida maatriksi \mathbf{X}^H pseudopöördmaatriks \mathbf{X}^+ . See oleks maatriks, mis etendaks tavalise pöördmaatriksiga sarnast rolli vähimruutude võrrandisüsteemide lahendamisel, kus tundmatute arv üldiselt ei võrdu võrrandite arvuga. Defineerime pseudopöördmaatriksi kui

$$\mathbf{X}^+ = \mathbf{V} \begin{bmatrix} \Sigma^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{U}^H, \quad \Sigma^{-1} = \text{diag}(\sigma_1^{-1}, \sigma_2^{-1}, \dots, \sigma_L^{-1}) \quad (6.46)$$

Lahendatav võrrandisüsteem võib olla kas alamääratud või ülemääratud. Mõlemal juhul avaldub lahend

$$\hat{\mathbf{w}} = \mathbf{X}^+ \mathbf{d} = \mathbf{V} \begin{bmatrix} \Sigma^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{U}^H \mathbf{d} \quad (6.47)$$

Vahe on selles, et ülemääratud võrrandisüsteemi (võrrandeid rohkem kui tundmatuid) pseudopöördmaatriks on

$$\mathbf{X}^+ = (\mathbf{X}\mathbf{X}^H)^{-1} \mathbf{X} \quad (6.48)$$

ning alamääratud võrrandisüsteemi (võrrandeid vähem kui tundmatuid) pseudopöördmaatriks on

$$\mathbf{X}^+ = \mathbf{X}(\mathbf{X}^H \mathbf{X})^{-1} \quad (6.49)$$

Kuna mõlemal juhul kehtib (6.46), siis ei ole enne võrrandisüsteemi lahendamisele asumist vaja teada, mis süsteemiga on tegemist. Kui lahendeid on rohkem kui üks, saame vähimruutude meetodit kasutades tulemuseks vähima eukleidese normiga lahendi.

Pseudopöördmaatriks üldistab pöördmaatriksi mõiste täisnurksetele maatriksitele. Saab näidata, et kehtivad nn Penrousi tingimused [7]

$$\begin{aligned}
\mathbf{X}^H \mathbf{X}^+ \mathbf{X}^H &= \mathbf{X}^H \\
\mathbf{X}^+ \mathbf{X}^H \mathbf{X}^+ &= \mathbf{X}^+ \\
(\mathbf{X}^H \mathbf{X}^+)^H &= \mathbf{X}^H \mathbf{X}^+ \\
(\mathbf{X}^+ \mathbf{X}^H)^H &= \mathbf{X}^+ \mathbf{X}^H
\end{aligned} \tag{6.50}$$

mis veelgi rõhutavad tavalise ruutmatriksi pöördmatriksi ja pseudopöördmatriksi sarnasust. Juhul kui \mathbf{X}^H on mittesingulaarne ruutmatriks, langeb pseudopöördmatriks kokku tavalise pöördmatriksiga

$$\mathbf{X}^+ = \mathbf{X}^{-H} \tag{6.51}$$

Näide

Olgu meil vaja lahendada võrrandisüsteem $\mathbf{X}^H \hat{\mathbf{w}} = \mathbf{d}$, kus

$$\mathbf{X}^H = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 0 & 0 \\ -1 & -2 & 1 \\ 0 & 0 & 2 \end{bmatrix} \text{ ja } \mathbf{d} = \begin{bmatrix} 2 \\ 5 \\ -2 \\ 3 \end{bmatrix}.$$

Esimesel pilgul tundub, et võrrandisüsteem on ülemääratud, kuna meil on neli võrrandit ja soovime leida kolme tundmatut. Lähemal vaatlusel selgub aga, et toodud võrrandisüsteem on vastuoluline – ei ole võimalik leida vektorit $\hat{\mathbf{w}}$, mis matriksi \mathbf{X}^H teise, nullise, reaga korrutades annaks tulemuseks vektori \mathbf{d} teise elemendi, 5. Peale selle on esimene ja kolmas võrrand teineteisest lineaarselt sõltuvad (ühe saame teisest, võttes mõlemad pooled vastandmärgiga). Sellegipoolest on võimalik leida võrrandisüsteemi lahend, kasutades lahutust singulaarväärtuste järgi.

Matriksi \mathbf{X}^H lahutus singulaarväärtuste järgi on

$$\mathbf{X}^H = \mathbf{U} \mathbf{S} \mathbf{V}^H = \begin{bmatrix} -0.67 & -0.21 & 0 & -0.71 \\ 0 & 0 & 1 & 0 \\ 0.67 & 0.21 & 0 & -0.71 \\ 0.30 & -0.95 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3.59 & 0 & 0 \\ 0 & 1.76 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -0.38 & -0.75 & 0.54 \\ -0.24 & -0.49 & -0.84 \\ 0.89 & -0.45 & 0 \end{bmatrix},$$

Edasi leiame võrrandisüsteemi lahendi, kasutades valemit (6.47). Selleks leiame kõigepealt singulaarväärtuste diagonaalmaatriksi pöördmatriksi, milleks saame

$$\mathbf{S}^+ = \begin{bmatrix} 0.28 & 0 & 0 & 0 \\ 0 & 0.57 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Leiame nüüd samm-sammult võrrandisüsteemi lahendi

$$\mathbf{U}^H \mathbf{d} = \begin{bmatrix} -1.78 \\ -3.71 \\ 5.0 \\ 0 \end{bmatrix}$$

$$\mathbf{S}^+ \mathbf{U}^H \mathbf{d} = \begin{bmatrix} -0.49 \\ -2.11 \\ 0 \end{bmatrix}$$

$$\hat{\mathbf{w}} = \mathbf{V} \mathbf{S}^+ \mathbf{U}^H \mathbf{d} = \begin{bmatrix} 0.7 \\ 1.4 \\ 1.5 \end{bmatrix}$$

Kotrolliks asendame leitud koefitsientide vektori võrrandisse endasse

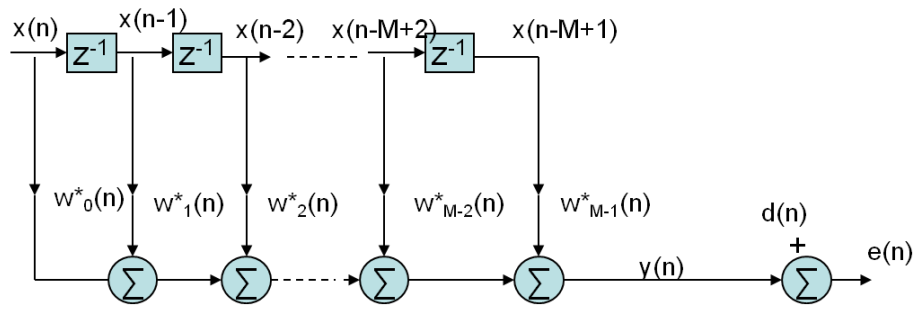
$$\mathbf{X}^H \hat{\mathbf{w}} = \begin{bmatrix} 2 \\ 0 \\ -2 \\ 3 \end{bmatrix}$$

Saadud vektori esimene, kolmas ja neljas element langevad kokku vektori \mathbf{d} vastavate elementidega. Teine element on 0, nagu võiski oodata, kuna maatriksi \mathbf{X}^H teine rida on nullvektor. Seega saime rahuldava lahendi.

Võrrandi lahendamine sellisel meetodil on numbriliselt stabiilne ning ei ole arvutuslikult liiga keerukas, kuna singulaarväärtuste järgi lahtuse leidmiseks on välja töötatud mitmeid kiireid algoritme nagu QR algoritm või Jacobi algoritm. Viimaste täpsem vaatlemine jääb küll väljapoole käesoleva õppevahendi piire. Huvitatud lugejale võib soovitada raamatut [1].

7. Rekursiivne vähimruutude meetod

Olles nüüd varustatud vähimruutude meetodiga, liigume edasi ühe olulise vähimruutude meetodil põhineva adaptiivse algoritmi – rekursiivse vähimruutude algoritmi (inglise keeles Recursive Least Squares) juurde. Nagu LMS algoritmigi puhul, kasutame ka rekursiivse vähimruutude algoritmi tähistamiseks edaspidi üldlevinud inglisekeelset lühendit RLS.



Joonis 48. Lõpliku impulsskajaga filter

Joonis 48 kujutab meile juba tuttavat lõpliku impulsskarakteristikuga filtrit. RLS algoritm alustab oma tööd mingist algolekust, tavaliselt nullisest koefitsientide vektorist $\mathbf{w}(0) = \mathbf{0}$ ning valib filtri koefitsiendid selliselt, et igal sammul minimeeritaks kriteeriumifunktsiooni, mis sõltub ajahetkeks n kogutud andmete hulgast ja seega muutub samm-sammult n -i kasvades

$$\sum_{i=1}^n \lambda^{n-i} |e(i)|^2, \quad 0 < \lambda < 1 \quad (7.1)$$

Näeme, et kriteeriumifunktsioon on eksponentsiaalselt kaalutud veasignaali diskreetide ruutude summa. Koefitsienti λ kutsutakse algoritmi unustamisfaktoriks ja selle väärtus valitakse ühest pisut väiksem. Unustamisfaktori ülesandeks on hoida saadavat algoritmi „elus“ nii, et ta unustaks aeglaselt vanad andmed ja oleks vastuvõtlik uutele. Ilma sellise unustamiseta ei suudaks algoritm mõni aeg hiljem enam muutunud oludega kohaneda.

Nagu õppevahendi eelmisteski osadeski, oleme kasutanud tähistusi

$$e(n) = d(n) - y(n) = d(n) - \mathbf{w}^H(n) \mathbf{x}(n) \quad (7.2)$$

kus sisendsignaali ja filtri kaalude vektorid ajahetkel n on defineeritud kui

$$\begin{aligned} \mathbf{x}(n) &= [x(n) \quad x(n-1) \quad \dots \quad x(n-M+1)]^T \\ \mathbf{w}(n) &= [w_0(n) \quad w_1(n) \quad \dots \quad w_{M-1}(n)]^T \end{aligned} \quad (7.3)$$

Nagu ikka, on kaalude vektor, mille korral veasignaali saavutab miinumumi, määratud normaalkõrranditega

$$\Phi(n) \mathbf{w}(n) = \mathbf{z}(n) \quad (7.4)$$

kusjuures antud juhul on sisendsignaali autokorrelatsioonimaatriks ning sisendsignaali ja soovitud signaali ristkorrelatsiooni vektor määratud kui

$$\Phi(n) = \sum_{i=1}^n \lambda^{n-i} \mathbf{x}(i) \mathbf{x}^H(i), \quad \mathbf{z}(n) = \sum_{i=1}^n \lambda^{n-i} \mathbf{x}(i) d^*(i) \quad (7.5)$$

Tuleb tähele panna, et ülaltoodud definitsioonid erinevad vähimruutude meetodi vastavatest definitsioonidest (6.13) ja (6.14) selle poolest, et nad sisaldavad unustamisfaktorit λ . Samuti eeldavad toodud avaldised, et signaalid on enne ajahetke $i = 1$ nullised.

Eraldades liikmed, mis vastavad ajahetkele n , saame rekursiivsed avaldised

$$\begin{aligned}\Phi(n) &= \lambda \left[\sum_{i=1}^{n-1} \lambda^{n-1-i} \mathbf{x}(i) \mathbf{x}^H(i) \right] + \mathbf{x}(n) \mathbf{x}^H(n) = \lambda \Phi(n-1) + \mathbf{x}(n) \mathbf{x}^H(n) \\ \mathbf{z}(n) &= \lambda \left[\sum_{i=1}^{n-1} \lambda^{n-1-i} \mathbf{x}(i) d^*(i) \right] + \mathbf{x}(n) d^*(n) = \lambda \mathbf{z}(n-1) + \mathbf{x}(n) d^*(n)\end{aligned}\tag{7.6}$$

mis seovad sisendsignaali autokorrelatsioonimaatriksi ning sisendsignaali ja soovitud signaali vahelise ristkorrelatsiooni vektori väärtused antud ajahetkel n vastavate väärtustega eelneval ajahetkel $n - 1$. Normaalkõrrandite lahend (6.8) sisaldab aga korrelatsioonimaatriksi asemel hoopis tema pöördmaatriksit ning seega sooviksime me leida avaldise (7.6) asemel rekursiivset valemit korrelatsioonimaatriksi pöördmaatriksi tarvis. Selleks kasutame matemaatikast tuntud tulemust, mis matemaatika-alases kirjanduses kannab Sherman-Morrison-Woodbury valemi nime [1], signaalitöötlejate hulgas on aga laialdaselt tuntud maatriksi pööramise lemma nime all.

Maatriksi pööramise lemma

Olgu \mathbf{A} ja \mathbf{B} kaks positiivselt määratud $M \times M$ maatriksit, \mathbf{D} positiivselt määratud $N \times N$ maatriks ja \mathbf{C} suvaline $M \times N$ maatriks. Ning kehtigu nimetatud maatriksite vahel seos

$$\mathbf{A} = \mathbf{B}^{-1} + \mathbf{C} \mathbf{D}^{-1} \mathbf{C}^H\tag{7.7}$$

Siis maatriksi \mathbf{A} pöördmaatriks avaldub kujul

$$\mathbf{A}^{-1} = \mathbf{B} - \mathbf{B} \mathbf{C} (\mathbf{D} + \mathbf{C}^H \mathbf{B} \mathbf{C})^{-1} \mathbf{C}^H \mathbf{B}\tag{7.8}$$

Maatriksi pööramise lemmat on lihtne tõestada, korrutades avaldise (7.7) ja (7.8) paremad pooled ning veendudes, et tulemuseks on ühikmaatriks.

Kasutame nüüd maatriksi pööramise lemmat korrelatsioonimaatriksi pöördmaatriksi leidmiseks avaldisest (7.6). Selleks valime valemities (7.7) ja (7.8) esinevad maatriksid kui

$$\mathbf{A} = \Phi(n), \quad \mathbf{B}^{-1} = \lambda \Phi(n-1), \quad \mathbf{C} = \mathbf{x}(n), \quad \mathbf{D} = 1\tag{7.9}$$

Taoline valik annab $\Phi(n)$ pöördmaatriksiks

$$\Phi^{-1}(n) = \lambda^{-1} \Phi^{-1}(n-1) - \frac{\lambda^{-2} \Phi^{-1}(n-1) \mathbf{x}(n) \mathbf{x}^H(n) \Phi^{-1}(n-1)}{1 + \lambda^{-1} \mathbf{x}^H(n) \Phi^{-1}(n-1) \mathbf{x}(n)}\tag{7.10}$$

Paneme tähele, et ülaltoodud avaldise teiseks liidetavaks oleva murru lugeja on maatriks, nimetaja aga skalaar. Edasi viime kirjepildi lihtsustamiseks sisse järgmised tähistused:

$$\mathbf{P}(n) = \Phi^{-1}(n)\tag{7.11}$$

ning

$$\mathbf{k}(n) = \frac{\lambda^{-1} \mathbf{P}(n-1) \mathbf{x}(n)}{1 + \lambda^{-1} \mathbf{x}^H(n) \mathbf{P}(n-1) \mathbf{x}(n)} \quad (7.12)$$

Maatriks $\mathbf{P}(n)$ on sisendsignaali korrelatsioonimaatriksi pöördmaatriks ajahetkel n ning vektor $\mathbf{k}(n)$ on osa avaldise (7.10) teisest liidetavast, mida kutsume võimenduse vektoriks (sellise nimetuse tagamaa selgub veidi hiljem). Kasutades niisuguseid tähistusi, võime avaldise (7.10) kirja panna ekvivalenttsel kujul

$$\mathbf{P}(n) = \lambda^{-1} \mathbf{P}(n-1) - \lambda^{-1} \mathbf{k}(n) \mathbf{x}^H(n) \mathbf{P}(n-1) \quad (7.13)$$

Edasi vaatleme lähemalt võimenduse vektorit $\mathbf{k}(n)$. Alustame $\mathbf{k}(n)$ definitsioonist valemis (7.12) ja korrutame avaldise (7.12) mõlemad pooled läbi parema poole nimetajaga, mis on skalaar. Seejärel viime liikme $\lambda^{-1} \mathbf{k}(n) \mathbf{x}^H(n) \mathbf{P}(n-1) \mathbf{x}(n)$ võrrandi paremale poole. Lõpuks viime sisendsignaali vektori ajahetkel n , $\mathbf{x}(n)$, mis nüüd esineb parema poole mõlemas liidetavas, sulgude taha, saades tulemuseks

$$\begin{aligned} \mathbf{k}(n) &= \lambda^{-1} \mathbf{P}(n-1) \mathbf{x}(n) - \lambda^{-1} \mathbf{k}(n) \mathbf{x}^H(n) \mathbf{P}(n-1) \mathbf{x}(n) \\ &= (\lambda^{-1} \mathbf{P}(n-1) - \lambda^{-1} \mathbf{k}(n) \mathbf{x}^H(n) \mathbf{P}(n-1)) \mathbf{x}(n) \end{aligned} \quad (7.14)$$

Sulgude sisse jäänud avaldis on aga valemi (7.13) põhjal võrdne sisendsignaali korrelatsioonimaatriksi pöördmaatriksiga ajahetkel n , seega võime (7.14) ümber kirjutada kujul

$$\mathbf{k}(n) = \mathbf{P}(n) \mathbf{x}(n) = \Phi^{-1}(n) \mathbf{x}(n) \quad (7.15)$$

Saime tulemuseks, et võimenduse vektor avaldub kui sisendsignaali korrelatsioonimaatriksi pöördmaatriks korda sisendsignaali vektor. Edasi vaatleme kaalude vektorit, mis avaldub normaalkorrannite (7.4) lahendina

$$\hat{\mathbf{w}}(n) = \Phi^{-1}(n) \mathbf{z}(n) = \mathbf{P}(n) \mathbf{z}(n) = \lambda \mathbf{P}(n) \mathbf{z}(n-1) + \mathbf{P}(n) \mathbf{x}(n) d^*(n) \quad (7.16)$$

Viimane võrdus ülaltoodud avaldises tuleneb ristkorrelatsiooni vektori rekursioonist (7.6).

Seejärel, asendades sisendsignaali korrelatsioonimaatriksi pöördmaatriksi rekursiooni (7.13) avaldise (7.16) esimesse liidetavasse, saame

$$\hat{\mathbf{w}}(n) = \mathbf{P}(n-1) \mathbf{z}(n-1) - \mathbf{k}(n) \mathbf{x}^H(n) \mathbf{P}(n-1) \mathbf{z}(n-1) + \mathbf{P}(n) \mathbf{x}(n) d^*(n) \quad (7.17)$$

Kuid valemi (7.16) põhjal on $\mathbf{P}(n-1) \mathbf{z}(n-1)$ võrdne kaalude vektori hinnanguga ajahetkel $n-1$. Seega saame kaalude rekursiooniks

$$\hat{\mathbf{w}}(n) = \hat{\mathbf{w}}(n-1) - \mathbf{k}(n) \mathbf{x}^H(n) \hat{\mathbf{w}}(n-1) + \mathbf{P}(n) \mathbf{x}(n) d^*(n) \quad (7.18)$$

Avaldisest (7.15) teame aga, et sisendsignaali korrelatsioonimaatriksi pöördmaatriks ajahetkel n korda sisendsignaali vektor ajahetkel n on võrdne võimenduse vektoriga ajahetkel n . Kasutades seda fakti eelneva avaldise viimases liidetavas ning tuues võimenduse vektori sulgude ette, saame

$$\hat{\mathbf{w}}(n) = \hat{\mathbf{w}}(n-1) + \mathbf{k}(n) [d^*(n) - \mathbf{x}^H(n) \hat{\mathbf{w}}(n-1)] \quad (7.19)$$

Sulgudes olev avaldis kujutab endast hinnangu viga, mille saame, kui kasutame ajahetkel $n-1$ leitud kaalude vektorit ning sisendsignaali vektorit ajahetkel n

$$\xi(n) = d(n) - \mathbf{x}^H(n)\mathbf{w}^*(n-1) = d(n) - \mathbf{w}^H(n-1)\mathbf{x}(n) \quad (7.20)$$

Seda viga kutsume *a priori* veaks, vastandina *a posteriori* hindamisveale

$$e(n) = d(n) - \mathbf{w}^H(n)\mathbf{x}(n) \quad (7.21)$$

mille leidmisel on kasutatud kaalude vektorit ja sisendsignaali vektorit ajahetkel n . *A priori* ja *a posteriori* vead ei ole üldjuhul omavahel võrdsed. Nendes tähistustes avaldub kaalude rekursioon kui

$$\hat{\mathbf{w}}(n) = \hat{\mathbf{w}}(n-1) + \mathbf{k}(n)\xi^*(n) \quad (7.22)$$

Siit saab ka selgeks, miks me kutsume vektorit \mathbf{k} võimenduse vektoriks. Nimelt tagasisidestatakse läbi selle võimenduse *a priori* viga algoritmi leitud kaaludesse.

Avaldised (7.11), (7.19) ning (7.13) määravadki RLS algoritmi, mille nüüd kokkuvõtlikult esitame

RLS algoritm

Initsialiseerimine: $\mathbf{P}(0) = \delta^{-1}\mathbf{I}$, $\delta =$ väike positiivne konstant
 $\hat{\mathbf{w}}(0) = \mathbf{0}$

Igal ajahetkel $n = 1, 2, \dots$ leida $\mathbf{k}(n) = \frac{\lambda^{-1}\mathbf{P}(n-1)\mathbf{x}(n)}{1 + \lambda^{-1}\mathbf{x}^H(n)\mathbf{P}(n-1)\mathbf{x}(n)}$

$$\xi(n) = d(n) - \mathbf{w}^H(n-1)\mathbf{x}(n)$$

$$\hat{\mathbf{w}}(n) = \hat{\mathbf{w}}(n-1) + \mathbf{k}(n)\xi^*(n)$$

$$\mathbf{P}(n) = \lambda^{-1}\mathbf{P}(n-1) - \lambda^{-1}\mathbf{k}(n)\mathbf{x}^H(n)\mathbf{P}(n-1)$$

Paneme tähele, et kuigi me alustasime algoritmi tuletust *a posteriori* vea minimeerimisest valemis (7.1), kasutab RLS algoritm oma rekursioonis *a priori* viga.

RLS algoritmi initsialiseerimine

Kui meil ei ole mingit lisainfot, mis soovitaks midagi muud, valitakse esialgne kaalude vektor tavaliselt nulline $\mathbf{w}(0) = \mathbf{0}$.

Maatriks \mathbf{P} on vaja initsialiseerida selliselt, et korrelatsioonimaatriks Φ oleks positiivselt määratud. Selle kindlustamiseks initsialiseeritakse korrelatsioonimaatriks tavaliselt diagonaalmaatriksiks, mille kõik diagonaalelemendid on võrdsed ja positiivsed. Seega eeldatakse korrelatsioonimaatriksi algväärtuse leidmisel, et sisendsignaali on valge müra. See annab meile kergelt modifitseeritud korrelatsioonimaatriksi rekursiooni

$$\Phi(n) = \sum_{i=1}^n \lambda^{n-i} \mathbf{x}(i)\mathbf{x}^H(i) + \delta\lambda^n \mathbf{I} \quad (7.23)$$

Seega on sisendsignaali korrelatsioonimaatriksi ning tema pöördmaatriksi algväärtused

$$\Phi(0) = \delta \mathbf{I}, \quad \mathbf{P}(0) = \delta^{-1} \mathbf{I} \quad (7.24)$$

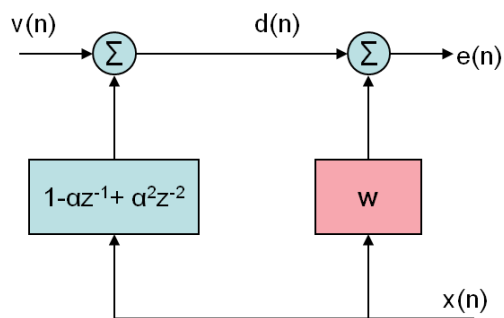
Ülaltoodud avaldised sõltuvad ainult ühest skalaarist δ , mis tuleb algoritmi initsialiseerimiseks valida. Praktilistele kogemustele toetudes soovitatakse parameeter δ valida väiksem kui $0.01\sigma_x^2$, kus σ_x^2 on sisendsignaali x võimsus. Tuleb märkida, et selline initsialiseerimine modifitseerib kergelt lahendatavat ülesannet, nii et RLS leiab tegelikult järgmise ülesande rekursiivse lahendi

$$\min_{\mathbf{w}(n)} \left[\delta \lambda^n \|\mathbf{w}(n)\|^2 + \sum_{i=1}^n \lambda^{n-i} |e(i)|^2 \right] \quad (7.25)$$

Tasub aga tähele panna, et kuna λ on ühest väiksem, väheneb andmete hulga n kasvades vahe kriteeriumite (7.25) ja (7.1) vahel kiiresti.

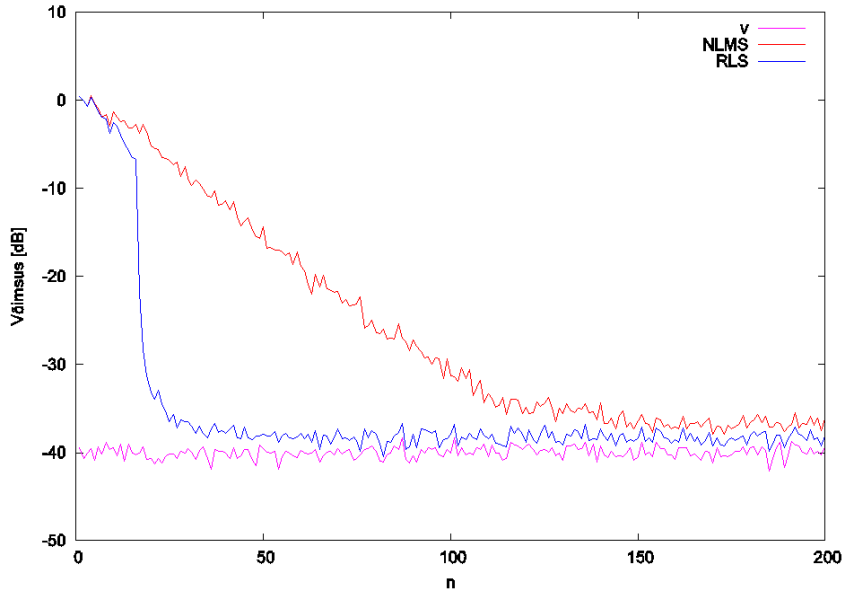
Näide

Vaatleme taas meile juba tuntud süsteemi identifitseerimise ülesannet. Joonis 49 kujutab uuritavat süsteemi, mille ülekandefunktsioon on $1 - \alpha z^{-1} + \alpha^2 z^{-2}$ kusjuures $\alpha = 0,2$.



Joonis 49. Uuritav süsteem

Selle süsteemi identifitseerimiseks kasutame 16 koefitsiendiga adaptiivset filtrit w . Sisendsignaaliks on valge müra võimsusega $\sigma_x^2 = 1$. Soovitud signaaliks on uuritava süsteemi väljundsignaal, millele lisandub sisendsignaalist statistiliselt sõltumatu valge müra võimsusega $\sigma_v^2 = 10^{-4}$.



Joonis 50. RLS ja NLMS algoritmide õppekõverad

Joonis 50 kujutab antud süsteemi identifitseerimise õppekõveraid, mis on leitud kasutades RLS ja normaliseeritud LMS algoritme. Võrdluseks on ära toodud ka mõõtemüra v võimsus. Kõik graafikud on saadud keskmistades üle saja realisatsiooni. NLMS algoritmi sammu suurus on valitud võrdseks ühega, tagamaks selle algoritmi kiireimat võimalikku koonduvust. RLS algoritmi unustamisfaktor on valitud $\lambda = 0,95$ ja korrelatsioonimaatriksi initialsiseerimiseks kasutatud parameeter $\delta = 10^{-3}$. Nagu jooniselt nähtub, koondub RLS algoritm oluliselt kiiremini kui NLMS. Tuleb aga märkida, et sellise koondumiskiiruse saavutamiseks kasutab RLS algoritm oluliselt rohkem arvutusi kui NLMS. ¶

Analüüs

Kasutades sõltumatuse teooria eeldusi ning eeldades, et RLS opereerib statsionaarses ümbruses, st optimaalne kaalude vektor \mathbf{w}_o ei muutu ajas, saab unustamisfaktori $\lambda = 1$ korral näidata, et RLS algoritmi kaalude vea korrelatsioonimaatriks avaldub kujul

$$\mathbf{K}(n) = E[(\hat{\mathbf{w}}(n) - \mathbf{w}_o)^H (\hat{\mathbf{w}}(n) - \mathbf{w}_o)] = \frac{\sigma^2}{n - M - 1} \mathbf{R}^{-1}, \quad n > M + 1 \quad (7.26)$$

Samadel eeldustel saab näidata, et RLS algoritmi õppimiskõver avaldub kui

$$J(n) = \sigma^2 + \text{tr}[\mathbf{R}\mathbf{K}(n-1)] = \sigma^2 + \frac{M\sigma^2}{n - M - 1}, \quad n > M + 1 \quad (7.27)$$

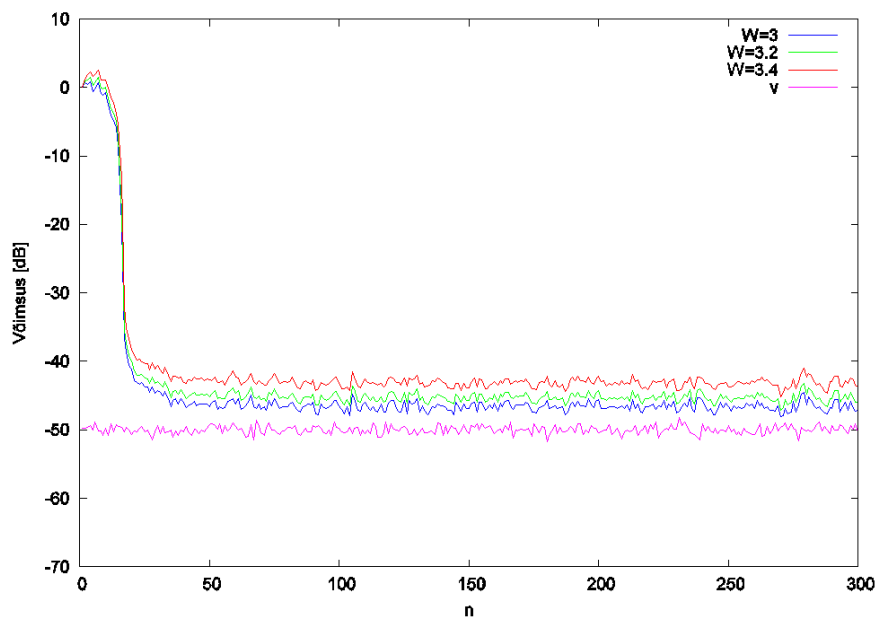
Suurus σ^2 on valemites (7.26) ja (7.27) võrdne vähimruutude mõttes optimaalse lahendi (6.1) vea e_o võimsusega. Toodud valemitest võime järeldada, et

1. RLS algoritm koondub tüüpiliselt $n = 2M$ iteratsiooniga, kui koondumise kriteeriumiks kasutada õppekõvera jõudmist 3dB lähedusse oma lõppväärtusest.

2. RLS algoritm koondub n -i lähenemisel lõpmatusse optimaalseks lahendiks $\lim_{n \rightarrow \infty} J(n) = \sigma^2$, ja seega n -i lähenemisel lõpmatusse lisaviga kaob. Võrdle LMS algoritmiga, mille lisaviga on esitatud valemiga (5.25).
3. Erinevalt kiireima languse meetodist ja LMS algoritmide perekonnast, ei sõltu RLS algoritmi koondumine sisendsignaali korrelatsioonimaatriksist ja tema omaväärtustest.

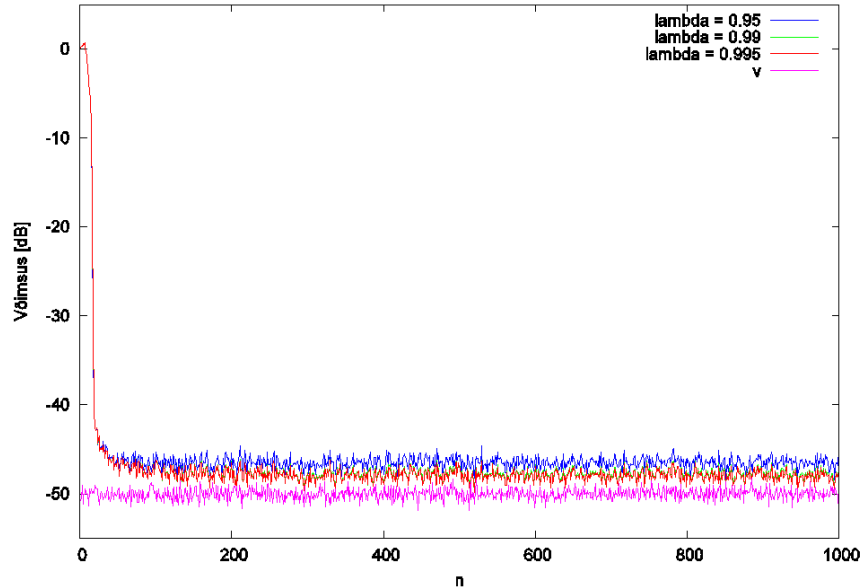
Näide

Vaatleme taas adaptiivset sidekanali ekvalaiseri näidet. Joonis 37 toob ära ekvalaiseri struktuurskeemi. Nagu ennegi iseloomustab kanalit ja allikat valem (5.32). Sisendsignaali võimsus on võrdne ühega ning kanali müra võimsus on 10^{-5} . Sidekanali pööramiseks on kasutatud filtrit pikkusega $M = 15$ ning soovitud signaali hilistumine on $D = 8$ diskreeti.



Joonis 51. RLS algoritmi koonduvuse sõltuvus kanali parameetritest

Joonis 51 kujutab RLS algoritmi õppekõveraid kolme erineva kanali parameetri W (3,0; 3,2 ja 3,4) korral. Nagu ennegi, on toodud graafikud saadud üle 100 realisatsiooni keskmistades. Vastavad õppekõverad LMS algoritmi jaoks toob ära Joonis 42. Võrreldes neid kaht joonist omavahel, näeme, et erinevalt LMS algoritmist, ei sõltu RLS algoritmi koondumiskiirus kanali parameetrist W ja seega ka mitte sisendsignaali korrelatsioonimaatriksi suurima ja vähima omaväärtuse suhtest (vaata Tabel 1). Tuleb tähele panna, et lõppvea sõltuvus kanali parameetrist ei tulene mitte RLS algoritmist, vaid sellest, et parameetri W suurenedes suureneb ka filtri pikkus, mis on vajalik kanali pööramiseks etteantud täpsusega. Teiste sõnadega jääb konstantse pikkusega pöördfiltri puhul lõppviga seda suuremaks, mida suurem on kanali parameeter W .



Joonis 52. RLS algoritmi koonduvuse sõltuvus unustamisfaktorist

Joonis 52 kujutab RLS algoritmi õppekõveraid kolme unustamisfaktori (0,95; 0,99 ja 0,995) korral. Nagu eespool, on kõverad saadud keskmistades üle 100 realisatsiooni. Kanali parameeter W on valitud võrdseks kolmega, mis annab suurima ja vähima sisendsignaali korrelatsioonimaatriksi omaväärtuse suhteks 8,54. Nagu jooniselt nähtub, ei sõltu RLS algoritmi koonduvuskiirus unustamisfaktorist, küll aga saavutab algoritm suurema unustamisfaktori korral väiksema lõppvea. Tuletame meelde, et unustamisfaktor peab olema väiksem või võrdne ühega. Lõppvea selline käitumine tuleb sellest, et suurema $\lambda \leq 1$ korral keskmistab algoritm efektiivselt pikema aja jooksul ning seega surutakse vaatlusmüra efektiivsemalt alla. Tuleb aga kohe märkida, et praktikas ei valita unustamisfaktorit tavaliselt siiski liiga ühe lähedale, andes sellega algoritmile võimaluse jälgida kiiremini süsteemi võimalikku muutumist ajas.

Libiseva aknaga RLS

Siiani vaatlesime RLS algoritmi, mis kasutas „vanade“ andmete eksponentsiaalset unustamist unustamisfaktori λ kaudu. On võimalik konstrueerida ka RLS algoritmi variant, mis kasutab eksponentsiaalse akna asemel täisnurkset konstantse pikkusega akent [1, 20]. Sellisel juhul unustatakse andmed teatud aja – akna pikkuse – möödumisel täielikult. Minimeeritav funktsioon oleks sel juhul (7.1) asemel

$$\sum_{i=n_1}^{n_2} |e(i)|^2 \quad (7.28)$$

Sellisel juhul oleksid meie ülesandele vastavad normaalvõrrandid

$$\begin{aligned} \Phi(n)\mathbf{w}(n) &= \mathbf{z}(n) \\ \Phi(n) &= \sum_{i=n_1}^{n_2} \mathbf{x}(i)\mathbf{x}^H(i), \quad \mathbf{z}(n) = \sum_{i=n_1}^{n_2} \mathbf{x}(i)d^*(i) \end{aligned} \quad (7.29)$$

ning algoritmi võime konstrueerida analoogselt eksponentsiaalse unustamisega RLS algoritmiga, vaid selle vahega, et nüüd ilmuvad sisendsignaali korrelatsioonimaatriksi ning sisendsignaali ja soovitud signaali ristkorrelatsiooni maatriksi avaldistesse liikmed, mis eemaldavad väljaspoole akent jäävad andmed. Sisendsignaali korrelatsioonimaatriksi rekursiooni võime siis kirjutada kujul

$$\begin{aligned}\Phi(n) &= \sum_{i=n_1}^{n_2-1} \mathbf{x}(i)\mathbf{x}^H(i) + \mathbf{x}(n_2)\mathbf{x}^H(n_2) - \mathbf{x}(n_1-1)\mathbf{x}^H(n_1-1) = \\ &= \Phi(n-1) + \mathbf{x}(n_2)\mathbf{x}^H(n_2) - \mathbf{x}(n_1-1)\mathbf{x}^H(n_1-1)\end{aligned}\quad (7.30)$$

ning sisendsignaali ja soovitud signaali korrelatsioonivektori rekursiooni kui

$$\begin{aligned}\mathbf{z}(n) &= \sum_{i=n_1}^{n_2-1} \mathbf{x}(i)d^*(i) + \mathbf{x}(n_2)d^*(n_2) - \mathbf{x}(n_1-1)d^*(n_1-1) = \\ &= \mathbf{z}(n-1) + \mathbf{x}(n_2)d^*(n_2) - \mathbf{x}(n_1-1)d^*(n_1-1)\end{aligned}\quad (7.31)$$

Sel viisil saadud algoritmide detaile vaata allikatest [1] või [20].

Kiired RLS algoritmid

RLS algoritim on arvutuslikult oluliselt keerukam, kui LMS perekonda kuuluvad algoritmid, vajades igal sammul $O(M^2)$ aritmeetilist operatsiooni, kus M on kasutatava filtri koefitsientide arv. Eksisteerivad ka kiired RLS algoritmid, mis vajavad oma tööks $O(M)$ aritmeetilist operatsiooni. Taolised algoritmid on leitavad näiteks allikatest [1 ja 11]. RLS algoritmi konkreetse versiooni valikul tuleb samas olla tähelepanelik [11], sest mitmetele neist on omane sisemine ebastabiilsus – algoritm küll koondub oma töö alguses, kuid mõne aja pärast võib tekkida järsk hajumine.

8. Algoritmid LMS ja RLS vahel

On terve hulk algoritme, mis oma keerukuselt ning omadustelt asuvad kusagil LMS ja RLS algoritmide vahepeal. Selles peatükis vaatlemegi neist mõnda, toetudes suures osas allikale [16]. Esimene selle perekonna algoritm on nn Newton-Raphsoni algoritm, mis põhineb Newtoni iteratiivsel meetodil võrrandi nullkohtade leidmiseks. Olgu meil vaja leida võrrandi $f(w) = 0$ lahend. Newtoni meetodi kohaselt võime lahendi leida, kasutades järgmist iteratiivset protseduuri [14]

$$w(i+1) = w(i) - \frac{f(w(i))}{f'(w(i))}\quad (8.1)$$

kus $f'(w)$ tähistab funktsiooni $f(w)$ tuletist w järgi.

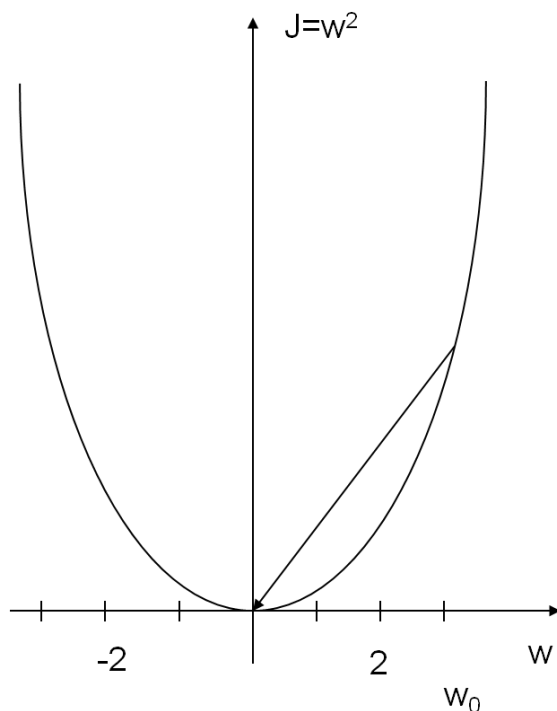
Adaptiivse filtreerimise ülesannetes ei soovi me otseselt võrrandit lahendada, kuid kriteeriumifunktsiooni minimeerimise ülesanne on sisuliselt kriteeriumifunktsiooni tuletise nullkohtade leidmine. Viimane ülesanne on aga võrrandi lahendamise ülesanne, mida on võimalik lahendada, kasutades rekursiooni (8.1).

$$\min_w J(w) \rightarrow \frac{\partial}{\partial w} J = 0\quad (8.2)$$

Rakendades Newtoni iteratiivset valemit ülaltoodud võrrandile, saame rekursiooni

$$w(i+1) = w(i) - \frac{\frac{\partial}{\partial w} J(w(i))}{\frac{\partial^2}{\partial w^2} J(w(i))} \quad (8.3)$$

Nagu kiireima languse meetodi puhul, alustame ka siin lihtsast ühemõõtmelisest ülesandest $J(w) = w^2$ ja vaatame, kuidas Newtoni meetod sel puhul töötab.



Joonis 53. Newtoni meetodi intuiitivne illustatsioon

Joonis 53 kujutabki kriteeriumifunktsiooni $J(w) = w^2$. Kriteeriumifunktsiooni esimene tuletis on $\frac{dJ(w)}{dw} = 2w$ ning teine tuletis $\frac{d^2J(w)}{dw^2} = 2$. Olgu meil alglähend $w_0 = 3$, siis $\left. \frac{dJ(w)}{dw} \right|_{w_0} = 2w_0 = 6$ ning $\left. \frac{d^2J(w)}{dw^2} \right|_{w_0} = 2$. Seega saame Newtoni meetodi kohaselt järgmiseks miinimumpunkti lähendiks $w_1 = w_0 - \frac{\left. \frac{dJ(w)}{dw} \right|_{w_0}}{\left. \frac{d^2J(w)}{dw^2} \right|_{w_0}} = w_0 - \frac{6}{2} = 0$, mis on ka kriteeriumi $J(w)$ tegelik miinimumpunkt.

Toodud näitest ilmneb, et kaasates teise tuletise, võimaldab Newtoni meetod leida mitte ainult kriteeriumifunktsiooni vähenemise suuna, nagu see oli kiireima languse meetodi rakendamisel, vaid ka sammu soovitava pikkuse, mis on vaja antud suunas astuda jõudmaks miinimumini.

Newton-Raphsoni meetod

Rakendame nüüd Newtoni meetodit M-mõõtmelise lõpliku impulsskarakteristikute filtri ülesandele. Sel juhul tuleb kriteeriumifunktsiooni esimene tuletis asendada tema gradiendiga ning teine tuletis

hessianiga. Hessian on $M \times M$ maatriks, mille elementideks on kriteeriumi teised tuletised kaalude vektori elementide järgi. Vastavalt Newton-Raphsoni meetodile astume igal iteratsioonil sammu suunas, mis on määratud hessiani pöördmaatriksi ja gradiendi korrutisega.

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu \left[\frac{\partial^2}{\partial \mathbf{w}^H \partial \mathbf{w}} J(n) \right]^{-1} \frac{\partial}{\partial \mathbf{w}^H} J(n) \quad (8.4)$$

Eelnevas avaldises sisalduv konstant μ võimaldab reguleerida algoritmi sammu pikkust. Tuletame meelde, et meid huvitava ülesande vea pind on ruutfunktsioon

$$J(n) = \delta_d^2 - \mathbf{p}^H \mathbf{w}(n) - \mathbf{w}^H(n) \mathbf{p} + \mathbf{w}^H(n) \mathbf{R} \mathbf{w}(n) \quad (8.5)$$

Võttes eeltoodust vajalikud tuletised, saame

$$\frac{\partial}{\partial \mathbf{w}^H} J(n) = -\mathbf{p} + \mathbf{R} \mathbf{w}(n) = -E[\mathbf{x}(n)e^*(n)] \quad (8.6)$$

ning

$$\frac{\partial}{\partial \mathbf{w}^H \partial \mathbf{w}} J(n) = \mathbf{R} \quad (8.7)$$

Asendades leitud tuletised valemisse (8.4), näeme, et Newton-Raphsoni meetodi rekursioon lõpliku impulsskarakteristikuga filtri koefitsientide leidmiseks on avaldatav kui

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu \mathbf{R}^{-1} [\mathbf{p} - \mathbf{R} \mathbf{w}(n)] = \mathbf{w}(n) + \mu \mathbf{R}^{-1} E[\mathbf{x}(n)e^*(n)] \quad (8.8)$$

Paneme tähele, et leitud avaldis sarnaneb kiireima languse meetodi rekursiooniga, kusjuures erinevuseks on sisendsignaali korrelatsioonimaatriksi pöördmaatriks, mis korrigeerib igal iteratsioonil astutava sammu suunda ja pikkust. Sarnaselt kiireima languse meetodiga on Newton-Raphsoni meetodi sellisel kujul rakendamiseks vaja teada üle ansambli keskmistatud statistikuid nagu sisendsignaali korrelatsioonimaatriks \mathbf{R} ja sisendsignaali ja soovitud signaali vaheline korrelatsioonivektor \mathbf{p} . Saamaks praktilistes situatsioonides sobivamat algoritmi, võime käituda sama moodi, nagu me tegime, kui tuletasime LMS algoritmi kiireima languse meetodist, st me võime asendada sisendsignaali vektori ja veasignaali kaaskompleksi korrutise matemaatilise ootuse tema hetkehinnanguga. Sisendsignaali korrelatsioonimaatriksi hinnanguna võime kasutada tema eksponentsiaalselt keskmistatud Toeplitzi aproksimatsiooni

$$T(n) = [t_{i-j}(n)] \quad (8.9)$$

$$t_k(n) = \lambda t_k(n-1) + x(n)x(n-k), \quad k = -M-1, \dots, M-1, \quad 0 < \lambda < 1$$

Nii toimides saame praktiliselt kasutatavaks algoritmiks

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu(n) \mathbf{T}^{-1}(n) \mathbf{x}(n) e^*(n), \quad (8.10)$$

Sammu suurus μ võib olla kas analoogselt LMS algoritmi põhikujuga konstant või võib kasutada ka ajas muutuvat sammu nii, nagu seda teeb normaliseeritud LMS algoritm. Viimasel juhul valitakse sammuks tihti

$$\mu(n) = \frac{1}{\mathbf{x}^H(n)\mathbf{T}^{-1}(n)\mathbf{x}(n)} \quad (8.11)$$

Afiinse projektsiooni algoritm

Vaatleme uuesti Newton-Rhaponi algoritmi. Rekursioon valemis (8.8) kasutab keskmistamist üle ansambli. Saamaks tema analoogi vähimruutude meetodi tähistustes, asendame sisendsignaali korrelatsioonimaatriksi vastava ajas keskmistatud korrelatsioonimaatriksiga valemist (6.13). Teguri $E[\mathbf{x}(n)\mathbf{e}^*(n)]$ ajas keskmistatud vasteks on analoogselt valemiga (6.14) avaldis $\mathbf{X}(n)\mathbf{e}^*(n)$, kus ülemine indeks * tähistab vektori kõigi elementide kaaskompleksi võtmist ilma vektorit transponeerimata. Rakendades eelöeldut, saame kaalude rekursioonile anda kuju

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu(\mathbf{X}(n)\mathbf{X}^H(n))^{-1}\mathbf{X}(n)\mathbf{e}^*(n) \quad (8.12)$$

Nüüd paneme tähele, et vastavalt valemile (6.48) kujutab tegur $(\mathbf{X}(n)\mathbf{X}^H(n))^{-1}\mathbf{X}(n)$ endast maatriksi $\mathbf{X}^H(n)$ pseudopöördmaatriksit $\mathbf{X}^+(n)$. Seega võime rekursiooni (8.12) kirja panna ekvivalentsel kujul

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu\mathbf{X}^+(n)\mathbf{e}^*(n) \quad (8.13)$$

Edasi näeme, et meil on võimalik valida, kui pikalt peaks algoritm signaale keskmistama, st millises suhtes peaksid olema tundmatute arv M ja ajaliseks keskmistamiseks kasutatud võrrandite arv L . Valik $L > M$ annab ülemääratud võrrandisüsteemi ja algoritm peab igal iteratsioonil pöörama $M \times M$ korrelatsioonimaatriksi, nagu seda teeb ka RLS algoritm. Teine võimalik valik on $L < M$, mis aga annab meile alamääratud võrrandisüsteemi. Kui kasutame nüüd alamääratud süsteemile vastavat pseudopöördmaatriksi määratlust valemist (6.49), võime rekursiooni (8.13) kirja panna kujul

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu\mathbf{X}(n)(\mathbf{X}^H(n)\mathbf{X}(n))^{-1}\mathbf{e}^*(n) \quad (8.14)$$

Paneme tähele, et erinevalt varasematest algoritmidest esineb valemis (8.14) veasignaali mitte skalaari vaid vektorina. Seega on meil enne kaalude rekursiooni vaja leida algoritmi igal sammul ka veasignaali vektor

$$\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{w}^H(n)\mathbf{X}(n) \quad (8.15)$$

Valemid (8.15) ja (8.14) koos moodustavadki afiinse projektsiooni algoritmi.

Huvitav on märkida, et valides eeltoodus võrrandite arvu L võrdseks ühega, taandub afiinse projektsiooni algoritm normaliseeritud LMS algoritmiks, mille toome võrdluse hõlbustamiseks uuesti välja

$$e(n) = d(n) - \mathbf{w}^H(n)\mathbf{x}(n)$$

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu}{\|\mathbf{x}(n)\|^2} \mathbf{x}(n)e^*(n) \quad (8.16)$$

Tulemus on loogiline, kuna valides võrrandite arvu L võrdseks ühega, oleme loobunud keskmistamisest samamoodi, nagu me loobusime keskmistamisest LMS algoritmi tuletamisel. Afiinse projektsiooni algoritmi praktilisel rakendamisel on mõistlik valida L lähtuvalt sisendsignaali omadustest nii, et suurem osa sisendsignaali korrelatsioonifunktsiooni oluliselt nullist erinevaid elemente paikneks lõigul $[0, L-1]$. Afiinse projektsiooni algoritmi oluliseks eeliseks ongi see, et kui RLS algoritmi puhul oli pööratav vajava sisendsignaali korrelatsioonimaatriksi suurus määratud kasutatava mudeli kaalude arvuga M , sõltumata sisendsignaalist, siis afiinse projektsiooni algoritm annab meile võimaluse pöörata maatriksit, mille suurus L on jätud disaineri valida ja mille saab siduda sisendsignaali omadustega.

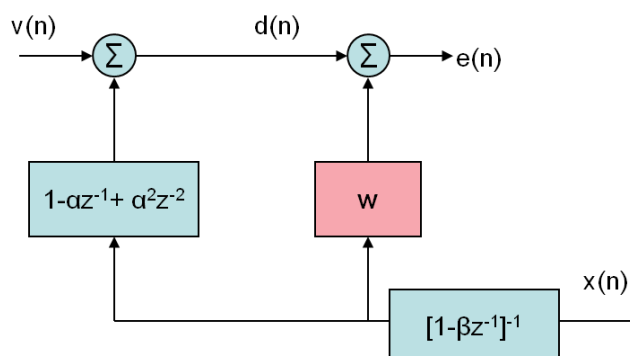
Nii nagu LMS algoritmi puhul, on ka afiinse projektsiooni algoritmi praktilisel rakendamisel mõistlik lisada nimetajasse väike konstant. Kuna nimetaja on antud juhul pööratav maatriks, liidame talle mitte skalaari vaid diagonaalmaatriksi $\sigma\mathbf{I}$, mis garanteerib pööratav maatriksi positiivse määratuse.

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu \mathbf{X}(n) (\mathbf{X}^H(n)\mathbf{X}(n) + \sigma\mathbf{I})^{-1} \mathbf{e}(n) \quad (8.17)$$

Nii toimides saame nn regulariseeritud afiinse projektsiooni algoritmi. Nagu RLS algoritmi puhulgi, eksisteerivad kiired algoritmid ka afiinse projektsiooni algoritmi tarvis. Nende algoritmide käsitus jääb aga väljapoole käesoleva õppevahendi raame ning huvitatud lugejal soovitame pöörduda allika [3] poole, kus on esitatud algoritm arvutusliku keerukusega $2M + 20L$ aritmeetilist operatsiooni iteratsiooni kohta.

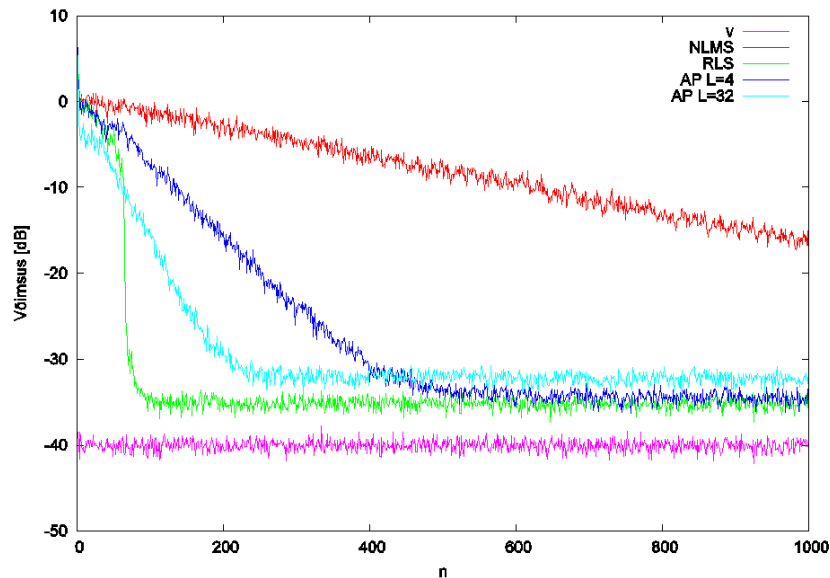
Näide

Joonis 54 kujutab näites uuritavat süsteemi, milleks on taas süsteemi identifitseerimise ülesanne



Joonis 54. Näites uuritav süsteem

Identifitseeritavaks süsteemiks on lõpliku impulsskarakteristikuga filter ülekandefunktsiooniga $1 - \alpha z^{-1} + \alpha^2 z^{-2}$, kusjuures $\alpha = 0,2$. Identifitseeritava süsteemi väljundsignaalile lisandub valge määtemüra $v(n)$ võimsusega $\sigma_v^2 = 10^{-4}$. Sisendsignaali on ühikulise võimsusega valge müra, mis on filtreeritud läbi lõpmatu impulsskarakteristikuga filtri ülekandefunktsiooniga $\frac{1}{1-\beta z^{-1}}$, kusjuures $\beta = 0,1$. Identifitseerimiseks kasutatav filter \mathbf{w} koosneb selles näites $M = 64$ koefitsiendist.



Joonis 55. Afiinse projektsiooni algoritmi õppekõverad

Joonis 55 kujutab afiinse projektsiooni algoritmi üle 100 realisatsiooni keskmistatud õppekõveraid parameetri L väärtuste 1 (normaliseeritud LMS), 4 ja 32 korral. Algoritmid on initsialiseeritud nullvektoriga, $\mathbf{w}(0) = \mathbf{0}$ ja sammu suurus μ on valitud võrdseks ühega. Nagu ikka on võrdluseks nagu ära toodud määtemüra võimsus, kuid ka RLS algoritmi õppekõver. Kuna sisendsignaal on selles näites tugevalt korreleeritud koondub normaliseeritud LMS suhteliselt aeglaselt, jõudmata graafikul näidatud 1000 iteratsiooniga väljakujunenud režiimi. RLS algoritm koondub oodatut kiirelt ning afiinse projektsiooni algoritmi koonduvuskiirus jääb normaliseeritud LMS ja RLS algoritmide vahele, kiirenedes parameetri L kasvades (kuna L -i kasvades suureneb korrelatsioonimaatriksi pööramisest tulenev sisendsignaali valgendusefekt).

9. Sagedusvallas toimivad adaptiivsed algoritmid

Üleminekut sagedusvalda kasutatakse kahel erineval põhjusel.

Esimeseks neist on soov kasutada diskreetsel Fourier' teisendusel (DFT) põhinevaid kiireid konvolutsiooni ja korrelatsiooni arvutamise algoritme. Teatavasti on lineaarse filtri väljundsignaali arvutamiseks vaja $O(M)$ aritmeetilist operatsiooni väljundsignaali diskreedi kohta, kus M on filtri koefitsientide arv. Kasutades DFT arvutamiseks kiire Fourier' teisenduse (FFT – Fast Fourier Transform) algoritme, võib selle keerukuse vähendada $O(\log_2 M)$ aritmeetilise operatsioonini [1, 9]. Sel teel võib ülesannetes, mis vajavad pikki adaptiivseid filtreid, saavutada olulise vajaliku arvutusvõimsuse kokkuhoiu. Efekti saavutamiseks peab filter olema piisava pikkusega, kuna FFT-l

põhinevate algoritmide keerukus on küll proportsionaalne logaritmiga filtri pikkusest, kuid proportsionaalsuse koefitsient on tavaliselt olulise suurusega.

Teiseks põhjuseks on lineaarsete ortogonaalsete teisenduste omadus jagada signaal ligikaudu sõltumatuteks osadeks selliselt, et tekib sisendsignaali valgendamise efekt ning saavutatakse kiirem koonduvus kui LMS algoritmi puhul.

Erinevalt selle õppevahendi ülejäänud osadest eeldame siin käsitluse lihtsustamiseks, et esinevad signaalid ei ole kompleksed, vaid omandavad ainult reaalseid väärtusi.

Alustame oma käsitlust algoritmidest, mis kasutavad diskreetset Fourier' teisendust adaptiivse algoritmi arvutusliku keerukuse vähendamiseks. Selleks märgime esmalt, et signaali $x(i)$ (pikkusega N diskreeti) diskreetne Fourier' teisendus on defineeritud kui

$$X(k) = \sum_{i=0}^{N-1} x(i) e^{-j2\pi ik/N}, \quad k = 0, \dots, N-1 \quad (9.1)$$

ja signaali $x(i)$ võib taastada tema diskreetsest Fourier' teisendusest $X(k)$, kasutades diskreetset Fourier' pöördteisendust

$$x(i) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) e^{j2\pi ik/N}, \quad i = 0, \dots, N-1 \quad (9.2)$$

Paneme tähele, et diskreetse Fourier' teisenduse (9.1) k -ndat elementi võime vaadelda kui signaalivektori ja vektori, mille i -ndaks elemendiks on kompleksne eksponent $e^{-j2\pi ik/N}$, skalaarkorrutist. Seega võime diskreetse Fourier' teisenduse panna kirja maatrikskujul

$$\begin{bmatrix} X(0) \\ X(1) \\ \vdots \\ X(N-1) \end{bmatrix} = \begin{bmatrix} e^{-j2\pi 0 \cdot 0/N} & e^{-j2\pi 1 \cdot 0/N} & \dots & e^{-j2\pi (N-1) \cdot 0/N} \\ e^{-j2\pi 0 \cdot 1/N} & e^{-j2\pi 1 \cdot 1/N} & \dots & e^{-j2\pi (N-1) \cdot 1/N} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-j2\pi 0 \cdot (N-1)/N} & e^{-j2\pi 1 \cdot (N-1)/N} & \dots & e^{-j2\pi (N-1) \cdot (N-1)/N} \end{bmatrix} \begin{bmatrix} x(0) \\ x(1) \\ \vdots \\ x(N-1) \end{bmatrix} \quad (9.3)$$

Tähistame eelnevas avaldises sisalduva diskreetse Fourier' teisenduse maatriksi kui

$$\mathbf{F} = \begin{bmatrix} e^{-j2\pi 0 \cdot 0/N} & e^{-j2\pi 1 \cdot 0/N} & \dots & e^{-j2\pi (N-1) \cdot 0/N} \\ e^{-j2\pi 0 \cdot 1/N} & e^{-j2\pi 1 \cdot 1/N} & \dots & e^{-j2\pi (N-1) \cdot 1/N} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-j2\pi 0 \cdot (N-1)/N} & e^{-j2\pi 1 \cdot (N-1)/N} & \dots & e^{-j2\pi (N-1) \cdot (N-1)/N} \end{bmatrix} \quad (9.4)$$

Analoogselt saame diskreetse Fourier' pöördteisenduse jaoks

$$\begin{bmatrix} x(0) \\ x(1) \\ \vdots \\ x(N-1) \end{bmatrix} = \frac{1}{N} \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & e^{j2\pi/N} & \dots & e^{j2\pi(N-1)/N} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{j2\pi(N-1)/N} & \dots & e^{j2\pi(N-1)(N-1)/N} \end{bmatrix} \begin{bmatrix} X(0) \\ X(1) \\ \vdots \\ X(N-1) \end{bmatrix} = \frac{1}{N} \mathbf{F}^H \begin{bmatrix} X(0) \\ X(1) \\ \vdots \\ X(N-1) \end{bmatrix} \quad (9.5)$$

On kerge veenduda, et maatriksid \mathbf{F} ja \mathbf{F}^H on unitaarsed ning \mathbf{F}^H/N on maatriksi \mathbf{F} pöördmaatriks

$$\mathbf{F}^{-1} = \frac{1}{N} \mathbf{F}^H \quad (9.6)$$

Edasi vaatleme tsirkulantmaatriksit. Tsirkulantmaatriks on maatriks, mis koosneb N erinevast elemendist c_0, \dots, c_{N-1} , mis on paigutatud vastavalt seaduspärasusele.

$$c_{k,l} = \begin{cases} c_{k-l}, & k \geq l \\ c_{n-l+1}, & k < l \end{cases} \quad (9.7)$$

Näiteks 4 x 4 tsirkulantmaatriksi oleks

$$\mathbf{C}_R = \begin{bmatrix} c_0 & c_1 & c_2 & c_3 \\ c_3 & c_0 & c_1 & c_2 \\ c_2 & c_3 & c_0 & c_1 \\ c_1 & c_2 & c_3 & c_0 \end{bmatrix} \quad (9.8)$$

Tsirkulantmaatriks pakub meile huvi ühe tema eriomaduse tõttu, nimelt diagonaliseerib DFT tsirkulantmaatriksi [8]. Teiste sõnadega on tsirkulantmaatriksi omavektorite maatriks DFT maatriks \mathbf{F} ning tema omaväärtused saame, kui arvutame diskreetse Fourier' teisenduse tsirkulantmaatriksi esimesest reast. Seega

$$\mathbf{C}_R = \frac{1}{\sqrt{n}} \mathbf{F} \mathbf{\Lambda} \frac{1}{\sqrt{n}} \mathbf{F}^H, \quad (9.9)$$

kusjuures tsirkulantmaatriksi \mathbf{C}_R omaväärtused λ_i , mis moodustavad diagonaalmaatriksi $\mathbf{\Lambda}$ peadiagonaali, on leitavad kui

$$\begin{bmatrix} \lambda_0 \\ \lambda_1 \\ \vdots \\ \lambda_{n-1} \end{bmatrix} = \mathbf{F} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{bmatrix} \quad (9.10)$$

Selle väite tõestamiseks kasutame diskreetse Fourier' teisenduse tsirkulaarse nihke omadust, mille kohaselt muutub signaali tsirkulaarne ajaline nihe Fourier' vallas korrutamiseks kompleksse eksponendiga. Seega järeljub seosest $x(n) \xleftrightarrow{DFT} X(k)$ seos

$x((n-m)_N) \xleftarrow{DFT} e^{-j2\pi km/N} X(k)$, kus $(\cdot)_N$ tähistab sulgudes oleva avaldise leidmist mooduliga N
 [9]. Näiteks $N = 4$ tsirkulantmaatriksi (9.8) jaoks saaksime

$$\begin{bmatrix} c_0 & c_1 & c_2 & c_3 \\ c_{(0-1)_N} & c_{1-1} & c_{2-1} & c_{3-1} \\ c_{(0-2)_N} & c_{(1-2)_N} & c_{2-2} & c_{3-2} \\ c_{(0-3)_N} & c_{(1-3)_N} & c_{(2-3)_N} & c_{3-3} \end{bmatrix} \begin{bmatrix} 1 \\ e^{-j2\pi k/N} \\ e^{-j2\pi 2k/N} \\ e^{-j2\pi 3k/N} \end{bmatrix} = \begin{bmatrix} X(k) \\ e^{-j2\pi k/N} X(k) \\ e^{-j2\pi 2k/N} X(k) \\ e^{-j2\pi 3k/N} X(k) \end{bmatrix} = X(k) \begin{bmatrix} 1 \\ e^{-j2\pi k/N} \\ e^{-j2\pi 2k/N} \\ e^{-j2\pi 3k/N} \end{bmatrix} \quad (9.11)$$

Loomulikult võime suurendada eelnevas avaldises sisalduva tsirkulantmaatriksi mõõdet N , suurendades samavõrra vektorite pikkusi, saades nii üldise valemi, mis kehtib suvalise N jaoks. Nüüd paneme tähele, et eelnevas avaldises sisalduvad eksponentidest koosnevad vektorid on diskreetse Fourier' teisenduse maatriksi \mathbf{F} veerud. Tähistades maatriksi \mathbf{F} veerud sümboliga \mathbf{f} , võime kirjutada

$$\mathbf{C}\mathbf{f}(k) = X(k)\mathbf{f}(k) \quad (9.12)$$

Valem (9.12) langeb aga kokku maatriksi \mathbf{C} omavektorite ja omaväärtuste definitsiooniga, kusjuures omaväärtusele $X(k)$ vastab omavektor $\mathbf{f}(k)$.

Edasi tuletame meelde plokk LMS algoritmi (5.51) reaalsete signaalide jaoks, mille võime maatrikskujul kirja panna kui

$$\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{X}(n)\mathbf{w}(n) \quad (9.13)$$

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu\mathbf{X}^T(n)\mathbf{e}(n)$$

kus $\mathbf{e}(n)$ on veasignaali vektor ajahetkel n ,

$$\mathbf{e}(n) = \begin{bmatrix} e(n-L+1) \\ \vdots \\ e(n-1) \\ e(n) \end{bmatrix}$$

$\mathbf{d}(n)$ on soovitud signaali vektor ajahetkel n ,

$$\mathbf{d}(n) = \begin{bmatrix} d(n-L+1) \\ \vdots \\ d(n-1) \\ d(n) \end{bmatrix}$$

$\mathbf{X}(n)$ on sisendsignaali maatriks ajahetkel n ,

$$\mathbf{X}(n) = \begin{bmatrix} x(n-L+1) & x(n-L) & \dots & x(n-M-L+2) \\ \vdots & \vdots & \dots & \dots \\ x(n-1) & x(n-2) & \ddots & x(n-M) \\ x(n) & x(n-1) & \dots & x(n-M+1) \end{bmatrix}$$

$\mathbf{w}(n)$ on adaptiivse filtri kaalude vektor ajahetkel n

$$\mathbf{w}(n) = \begin{bmatrix} w_0(n) \\ w_1(n) \\ \vdots \\ w_{M-1}(n) \end{bmatrix}$$

ning μ on algoritmi sammu suurus.

Edasi moodustame tsirkulantse sisendsignaali maatriksi, laiendades maatriksi \mathbf{X} sisendsignaali $x(n)$ diskreetidega selliselt, et tulemuseks oleks tsirkulantne maatriks ja esialgne sisendsignaali maatriks \mathbf{X} jääks tsirkulantmaatriksi alumisse vasakusse nurka.

$$\mathbf{X}_{T_s}(n) = \begin{bmatrix} x(n-M-L+1) & x(n-M-L) & x(n-M-L-1) & \dots & x(n-M-L+2) \\ x(n-M-L+2) & x(n-M-L+1) & x(n-M-L) & \dots & x(n-M-L+3) \\ \dots & \dots & \dots & \ddots & \dots \\ x(n-L+1) & x(n-L) & x(n-L-1) & \dots & x(n-L+2) \\ \dots & \dots & \dots & \ddots & \dots \\ x(n-1) & x(n-2) & x(n-3) & \dots & x(n) \\ x(n) & x(n-1) & x(n-2) & \dots & x(n-M-L+1) \end{bmatrix}$$

Esialgse sisendsignaali maatriksi \mathbf{X} võime saada maatriksist \mathbf{X}_{T_s} , korrutades teda nii vasakult kui paremalt maatriksitega, mis koosnevad nullmaatriksi ja ühikmaatriksi ühendusest.

$$\mathbf{X}(n) = [\mathbf{0}_{L \times M} \quad \mathbf{I}_L] \mathbf{X}_{T_s}(n) \begin{bmatrix} \mathbf{I}_M \\ \mathbf{0}_{L \times M} \end{bmatrix} \quad (9.14)$$

Alumised indeksid maatriksi tähistuse juures näitavad maatriksi dimensioone. Näiteks $L = M = 2$ korral annaks tsirkulantmaatriksi korrutamine ühikmaatriksi ja nullmaatriksi kombinatsiooniga vasakult

$$\begin{bmatrix} x(n-3) & x(n) & x(n-1) & x(n-2) \\ x(n-2) & x(n-3) & x(n) & x(n-1) \\ x(n-1) & x(n-2) & x(n-3) & x(n) \\ x(n) & x(n-1) & x(n-2) & x(n-3) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} x(n-3) & x(n) \\ x(n-2) & x(n-3) \\ x(n-1) & x(n-2) \\ x(n) & x(n-1) \end{bmatrix}$$

Selline operatsioon eraldab tsirkulantmaatriksi kaks esimest veergu. Sarnaselt eraldab ühikmaatriksi ja nullmaatriksi kombinatsiooniga korrutis paremalt tsirkulantmaatriksi kaks viimast rida

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x(n-3) & x(n) & x(n-1) & x(n-2) \\ x(n-2) & x(n-3) & x(n) & x(n-1) \\ x(n-1) & x(n-2) & x(n-3) & x(n) \\ x(n) & x(n-1) & x(n-2) & x(n-3) \end{bmatrix} = \begin{bmatrix} x(n-1) & x(n-2) & x(n-3) & x(n) \\ x(n) & x(n-1) & x(n-2) & x(n-3) \end{bmatrix}$$

Avaldises (9.14) oleme kasutanud mõlemat äsjakirjeldatud võtet eraldamaks tsirkulantmaatriksi alumist vasakpoolset kvadranti, kus asub meie esialgne sisendsignaali maatriks.

Plokk LMS algoritmi (9.13) jaoks vajame ka transponeeritud sisendsignaali maatriksit, mille saame valemist (9.14) kui

$$\mathbf{X}^T(n) = \begin{bmatrix} \mathbf{I}_M & \mathbf{0}_{M \times L} \end{bmatrix} \mathbf{X}_{Ts}^T(n) \begin{bmatrix} \mathbf{0}_{M \times L} \\ \mathbf{I}_L \end{bmatrix} \quad (9.15)$$

Kuna \mathbf{X}_{Ts} on reaalne maatriks, siis tema transpositsioon võrdne maatriksi \mathbf{X}_{Ts} Hermiti transpositsiooniga, seega, kasutades tsirkulantmaatriksi lahutust omaväärtuste ja omavektorite kaudu avaldises (9.9) ning jättes kompaktsema kirja pildi huvides kirjutamata aega väljendava indeksi n , saame

$$\mathbf{X}_{Ts} = \frac{1}{M+L} \mathbf{F} \Lambda^H \mathbf{F}^H = \frac{1}{M+L} \mathbf{F}^H \Lambda \mathbf{F} \quad (9.16)$$

Eelneva võrrandi parem pool on saadud pärast tavalise transpositsiooni ning seejärel Hermiti transpositsiooni leidmist. Kuna nimetatud kaks operatsiooni on reaalsete maatriksite puhul ekvivalentid, saame pärast kaht sellist operatsiooni maatriksi, mis on võrdne esialgse maatriksiga. Anaoloogiliselt leiame transponeeritud sisendsignaali tsirkulantse laiendi tarvis

$$\mathbf{X}_{Ts}^T = \frac{1}{M+L} \mathbf{F} \Lambda \mathbf{F}^H = \frac{1}{M+L} \mathbf{F}^H \Lambda^H \mathbf{F} \quad (9.17)$$

Kuna meie tsirkulantne maatriks omab mõõtmeid $(M+L) \times (M+L)$, siis on ka diskreetse Fourier' teisenduse maatriksi \mathbf{F} mõõtmeteks $(M+L) \times (M+L)$. Seega läheb meil algoritmi tööks vaja diskreetset Fourier' teisendust, mille pikkus on võrdne filtri koefitsientide arvu M ning signaali ploki pikkuse L summaga.

Edasi asendame sisendsignaali maatriksi avaldise

$$\mathbf{X} = \frac{1}{M+L} \begin{bmatrix} \mathbf{0}_{L \times M} & \mathbf{I}_L \end{bmatrix} \mathbf{F}^H \Lambda \mathbf{F} \begin{bmatrix} \mathbf{I}_M \\ \mathbf{0}_{L \times M} \end{bmatrix} \quad (9.18)$$

ja vastava transponeeritud sisendsignaali maatriksi

$$\mathbf{X}^T = \frac{1}{M+L} \begin{bmatrix} \mathbf{I}_M & \mathbf{0}_{M \times L} \end{bmatrix} \mathbf{F}^H \Lambda^H \mathbf{F} \begin{bmatrix} \mathbf{0}_{M \times L} \\ \mathbf{I}_L \end{bmatrix} \quad (9.19)$$

plokk LMS algoritmi võrranditesse (9.13), ning saame veavektori arvutamise võrrandiks

$$\mathbf{e}(n) = \mathbf{d}(n) - \frac{1}{M+L} \begin{bmatrix} \mathbf{0}_{L \times M} & \mathbf{I}_M \end{bmatrix} \mathbf{F}^H \Lambda \mathbf{F} \begin{bmatrix} \mathbf{I}_M \\ \mathbf{0}_{L \times M} \end{bmatrix} \mathbf{w}(n) \quad (9.20)$$

ning adaptiivse filtri kaalude rekursiooniks

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu \frac{1}{M+L} \begin{bmatrix} \mathbf{I}_M & \mathbf{0}_{M \times L} \end{bmatrix} \mathbf{F}^H \Lambda^H \mathbf{F} \begin{bmatrix} \mathbf{0}_{M \times L} \\ \mathbf{I}_L \end{bmatrix} \mathbf{e}(n) \quad (9.21)$$

Järgnevalt defineerime adaptiivse filtri kaalude vektori sagedusvallas kui diskreetse Fourier' teisenduse nullidega laiendatud ajavalla kaalude vektorist, nii nagu ta figureerib veavektori arvutamise võrrandis (9.20)

$$\mathbf{f}(n) = \mathbf{F} \begin{bmatrix} \mathbf{I}_M \\ \mathbf{0}_{L \times M} \end{bmatrix} \mathbf{w}(n) \quad (9.22)$$

$M \times 1$ ajavalla kaalude vektori $\mathbf{w}(n)$ korrutamisel $(L+M) \times M$ maatriksiga $\begin{bmatrix} \mathbf{I}_M \\ \mathbf{0}_{L \times M} \end{bmatrix}$ saame tulemuseks $(L+M) \times 1$ vektori, mille esimesed M elementi langevad kokku vektori $\mathbf{w}(n)$ elementidega ning viimased L elementi on nullised. Seetõttu võime sellist korrutamisoperatsiooni käsitleda kui kaalude vektori täiendamist nullidega. Kasutades tähistust (9.22), võime veavektori leidmise võrrandi (9.20) ümber kirjutada kujul

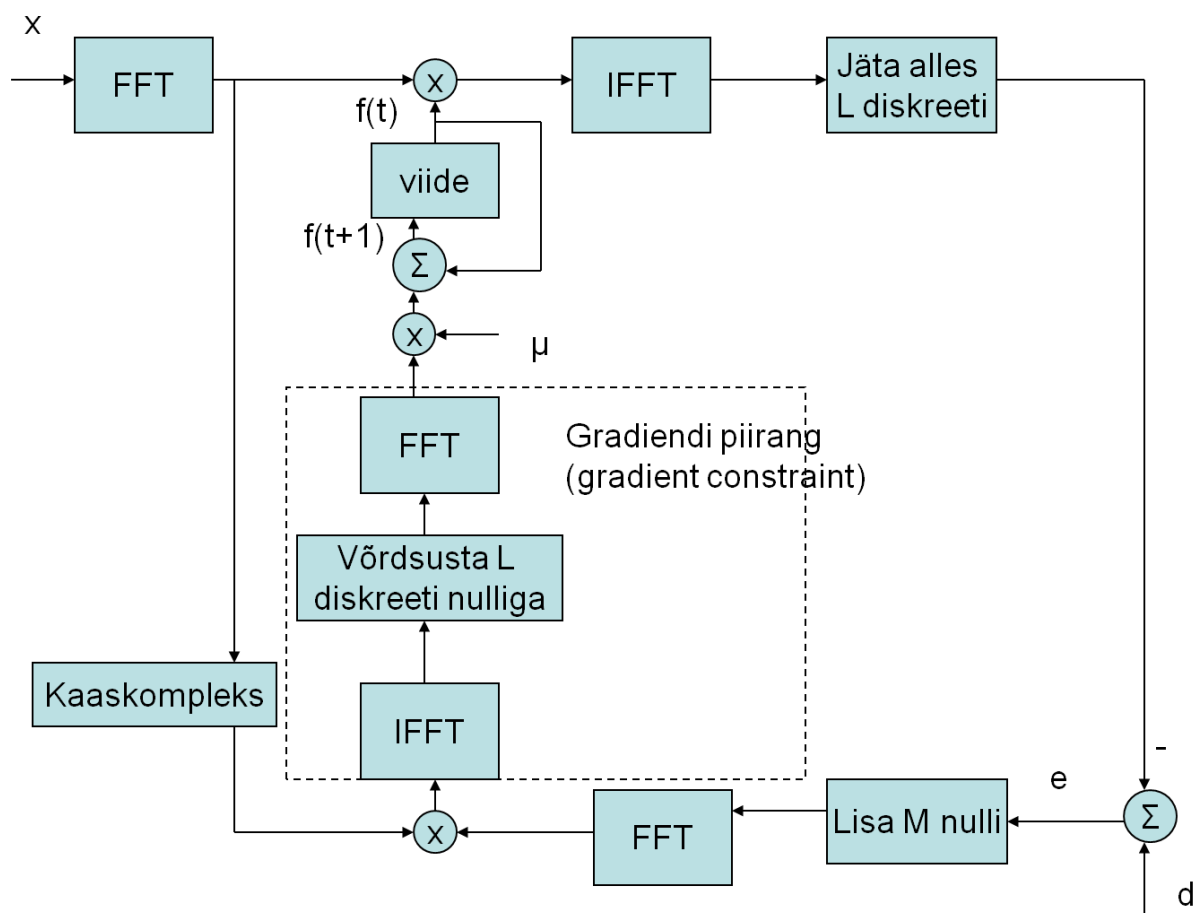
$$\mathbf{e}(n) = \mathbf{d}(n) - \frac{1}{M+L} \begin{bmatrix} \mathbf{0} & \mathbf{I} \end{bmatrix} \mathbf{F}^H \Lambda \mathbf{f}(n) \quad (9.23)$$

Kaalude rekursiooni (9.21) korrutame vasakult maatriksiga $\mathbf{F} \begin{bmatrix} \mathbf{I}_M \\ \mathbf{0}_{L \times M} \end{bmatrix}$ ning kasutame sagedusvalla kaalude tähistust valemist (9.22), saades tulemuseks sagedusvalla kaalude rekursiooni kujul

$$\mathbf{f}(n+1) = \mathbf{f}(n) + \mu \frac{1}{M+L} \mathbf{F} \begin{bmatrix} \mathbf{I}_M \\ \mathbf{0}_{L \times M} \end{bmatrix} \begin{bmatrix} \mathbf{I}_M & \mathbf{0}_{M \times L} \end{bmatrix} \mathbf{F}^H \Lambda^H \mathbf{F} \begin{bmatrix} \mathbf{0}_{M \times L} \\ \mathbf{I}_L \end{bmatrix} \mathbf{e}(n) \quad (9.24)$$

Valemid (9.23) ja (9.24) koos moodustavadki otsitud sagedusvalla adaptiivse algoritmi. Paneme tähele, et veavektori korrutamine $(L+M) \times L$ maatriksiga $\begin{bmatrix} \mathbf{0}_{M \times L} \\ \mathbf{I}_L \end{bmatrix}$ vasakult (nii nagu valemis (9.24) on tehtud $L \times 1$ vektoriga $\mathbf{e}(n)$) annab tulemuseks vektori, mida on tipust laiendatud M nulliga.

Sagedusvalla kaalude rekursioonis esinevat maatriksit $\mathbf{F} \begin{bmatrix} \mathbf{I}_M \\ \mathbf{0}_{L \times M} \end{bmatrix} \begin{bmatrix} \mathbf{I}_M & \mathbf{0}_{M \times L} \end{bmatrix} \mathbf{F}^H$ kutsutakse tihti gradiendi piiranguks. Vektori vasakult korrutamine sellise maatriksiga on ekvivalentne operatsiooniga, mille käigus arvutatakse kõigepealt vektori diskreetne Fourier' pöördteisendus, asendatakse tema L viimast elementi nullidega ning lõpuks leitakse tulemusest diskreetne Fourier' teisendus.



Joonis 56. Sagedusvalla adaptiivne algoritm

Joonis 56 kujutab äsjatuletatud sagedusvalla adaptiivset algoritmi graafiliselt. Algoritm leiab iga $L+M$ diskreeti pikkuse sisendplokki (millest M diskreeti kattuvad eelmises plokis sisaldunud diskreetidega) diskreetse Fourier' teisenduse. Leitud sisendsignaali diskreetne Fourier' teisendus korrutatakse sagedusvalla filtrikoefitsientidega $f(n)$ leidmaks filtri väljundsignaali diskreetset Fourier' teisendust ning seejärel arvutatakse tulemustest diskreetne Fourier' pöördteisendus, st viiakse filtri väljundsignaal ajavalda. Filtri väljundsignaalist jäetakse alles plokki pikkusele vastavad L diskreeti. Seejärel leitakse veasignaali $e(n)$, lahutades L diskreeti pikkusest soovitud signaali plokist $d(n)$ äsjaleitud filtri väljundsignaali plokki. Edasi lisatakse pikkusega L veasignaali plokile M nulli ning tulemustest arvutatakse diskreetne Fourier' teisendus, mis korrutatakse eelnevalt leitud sisendsignaali diskreetse Fourier' teisenduse kaaskompleksiga. Kõrvalmärkusena nendime, et selline operatsioon vastaks ajavallas signaalide korrelatsioonifunktsiooni arvutamisele. Edasi leiame diskreetse Fourier' pöördteisenduse, mille viimased L elementi võrdsustame nulliga ning leiame tulemustest uuesti diskreetse Fourier' teisenduse. Tulemi korrutame sammu suurusega μ ning liidame olemasolevatele adaptiivse filtri koefitsientidele, saamaks kaale, mis tulevad kasutusele algoritmi järgmisel sammul.

Märgime, et toodud adaptiivse filtreerimise tehnikal on mitmeid sarnasusi diskreetsete signaalide filtreerimisega diskreetset Fourier' teisendust kasutades, kus samuti tuleb õige tulemuse saamiseks signaalidele sobivalt nulle lisada ning osast signaalelementidest loobuda [1, 9].

Kuna tuletasime oma sagedusvalla adaptiivse algoritmi plokk LMS algoritmist, asendades arvutused ajavallas arvutustega sagedusvallas, siis on ka leitud algoritmi omadused väga sarnased plokk LMS algoritmi omadele. Tuleb välja, et sagedusvalla algoritmi omadusi on võimalik oluliselt parandada, normaliseerides vanadele kaaludele lisatavat liiget sisendsignaali võimsusspektri hinnanguga. Kasutades sellist normaliseerimist, saaksime uueks sagedusvalla kaalude rekursiooniks

$$\mathbf{f}(n+1) = \mathbf{f}(n) + \mu \frac{1}{M+L} \mathbf{F} \begin{bmatrix} \mathbf{I}_M \\ \mathbf{0}_{L \times M} \end{bmatrix} \begin{bmatrix} \mathbf{I}_M & \mathbf{0}_{M \times L} \end{bmatrix} \mathbf{F}^H \Lambda^H \mathbf{P}^{-1}(n) \mathbf{F} \begin{bmatrix} \mathbf{0}_{M \times L} \\ \mathbf{I}_L \end{bmatrix} \mathbf{e}(n) \quad (9.25)$$

kusjuures normaliseeriv maatriks \mathbf{P} oleks diagonaalmaatriks, mille peadiagonaalil asuvad sisendsignaali võimsusspektri hinnangu diskreedid

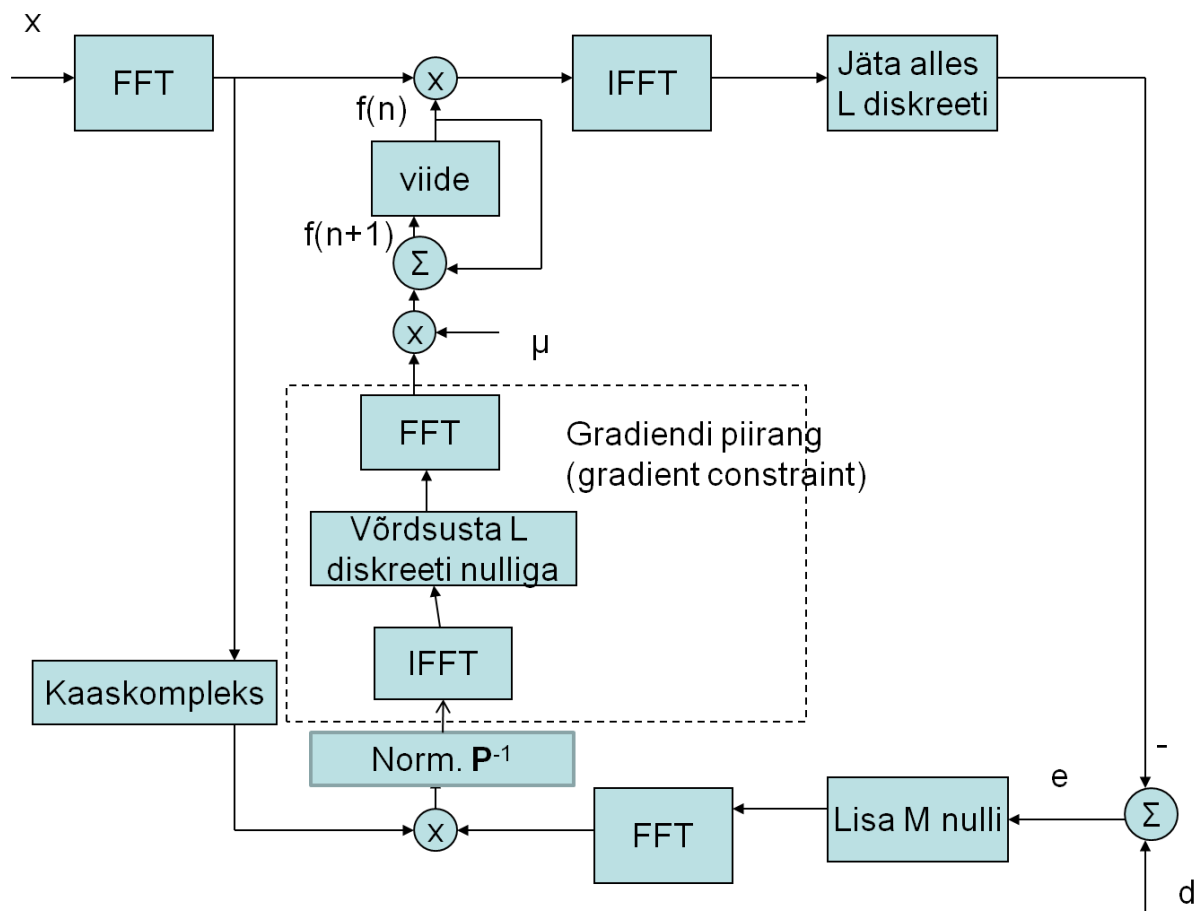
$$\mathbf{P}(n) = \text{diag}(P_0(n), P_1(n), \dots, P_{L+M-1}(n)) \quad (9.26)$$

ning need võib leida, keskmistades eksponentsiaalselt sisendsignaali diskreetse Fourier' teisenduse elementide mooduli ruute.

$$P_i(n) = \lambda P_i(n-1) + (1-\lambda) |X_i(n)|^2 \quad (9.27)$$

Eelolevas valemis esinev konstant λ tuleks valida vahemikus $0 < \lambda < 1$.

Paneme tähele, et selline normaliseerimine sarnaneb näiteks Newton-Raphsoni algoritmis eksisteeriva normalisatsiooniga. Vahe on selles, et kui Newton-Raphsoni algoritm (8.8) kasutas normalisatsiooniks sisendsignaali korrelatsioonimaatriksit, siis siin kasutame sama efekti saavutamiseks autokorrelatsioonifunktsiooni sagedusvalla ekvivalenti – võimsusspektrit.

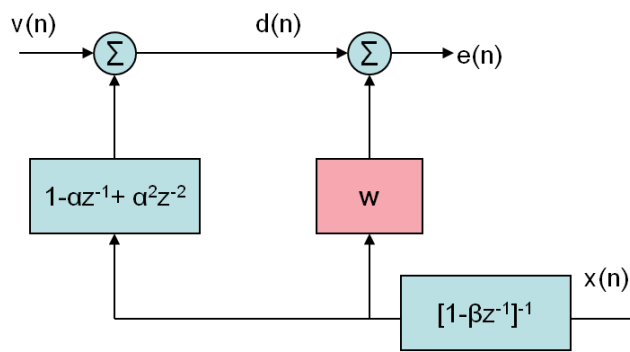


Joonis 57. Normaliseeritud sagedusvalla adaptiivne algoritm

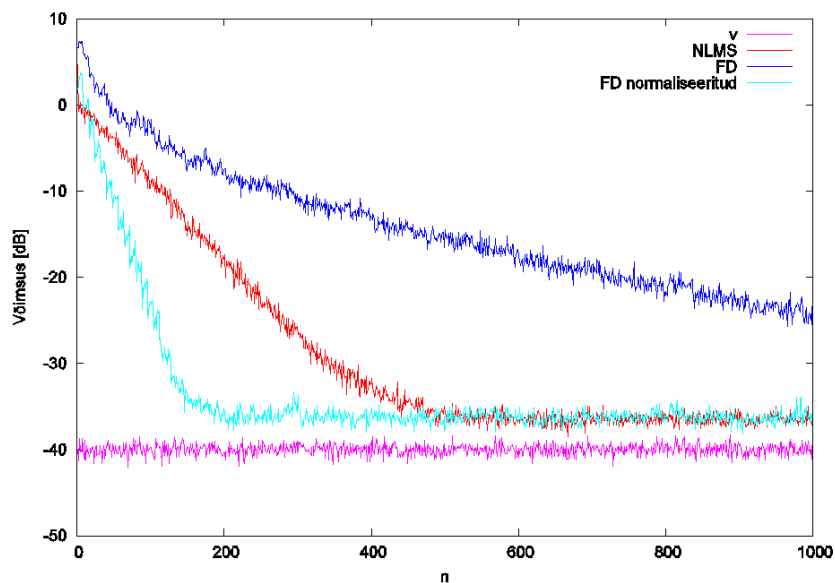
Joonis 57 kujutab normaliseeritud sagedusvalla adaptiivset algoritmi graafiliselt.

Näide

Vaatleme taas meile eelnevast tuttavat süsteemi identifitseerimise ülesannet. Joonis 58 kujutab näites kasutatavat ülesannet graafiliselt. Sisendsignaalina on kasutusel ühikulise võimsusega valge müra, mis on filtreeritud läbi filtri ülekandefunktsiooniga $\frac{1}{1-\beta z^{-1}}$, kusjuures $\beta = 0,1$. Identifitseeritava süsteemi ülekandefunktsioon on $1 - \alpha z^{-1} + \alpha^2 z^{-2}$, kusjuures $\alpha = 0,2$. Identifitseeritava süsteemi väljundsignaalile lisandub valge mõõtemüra võimsusega $\sigma_v^2 = 10^{-4}$. Identifitseerimiseks kasutatakse filter \mathbf{w} koosneb $M = 16$ lülist ning ploki pikkuseks on $L = 8$.



Joonis 58. Uuritav süsteem



Joonis 59. Õppekõverad

Joonis 59 kujutab meie normaliseerimata (9.24) ja normaliseeritud (9.25) sagedusvalla algoritmide õppekõveraid. Võrdluseks on toodud ka normaliseeritud LMS algoritmi õppekõver ning mõõtemüra võimsus. Kõik graafikud on keskmistatud üle 100 realisatsiooni. Sammu pikkus μ on valitud võrdseks ühega. Nagu graafikult näha, koondub normaliseerimata sagedusvalla algoritm selles situatsioonis aeglasemalt kui normaliseeritud LMS. See on tingitud algoritmi plokistruktuurist. Normaliseeritud sagedusvalla algoritm koondub see-eest oluliselt kiiremini kui ülejäänud kaks algoritmi.

Oluline on tähele panna, et diskreetse Fourier' teisenduse arvutamiseks on vajalik signaali plokk pikkusega L diskreeti, mis tuleb enne iga ploki töötlemist kokku koguda. Seega toob selline sagedusvalla algoritm sisse ajalise viite, mis on võrdne ploki pikkuse ehk L diskreediga. Taoline viide ei

kujuta paljudes rakendustes endast mingit probleemi, kuid ei ole aktsepteeritav mitmete siderakenduste, näiteks kajakõrvaldajate, korral.

Arvutuslik keerukus

$(M + L)$ -punktise diskreetse Fourier' teisenduse leidmiseks kulub juhul, kui arvutusteks kasutada mõnd kiire Fourier' teisenduse algoritmi [9] $O((M + L) \log_2(M + L))$ aritmeetilist operatsiooni. Siin oleme eeldanud, et $(M + L)$ on mingi kahe aste, kuna just sellise pikkusega signaalide jaoks eksisteerivad kõige efektiivsemad kiire Fourier' teisenduse algoritmid.

Algoritm vajab iga pikkusega L signaaliploki töötlemiseks kolme diskreetset Fourier' teisendust, kahte pöördteisendust, kahte pikkusega $(M + L)$ vektorite elemendikaupa korrutamist, ühte pikkusega L vektorite lahutamist ning lineaarse keerukusega normaliseerimisoperatsiooni (jagamist). Seega on algoritmi kogukeerukus $O((M + L) \log_2(M + L))$ aritmeetilist operatsiooni pikkusega L signaaliploki kohta või siis vastavalt $O\left(\frac{(M + L)}{L} \log_2(M + L)\right)$ aritmeetilist operatsiooni diskreedi kohta.

Praktikas valitakse sageli töödeldava signaaliploki pikkus L võrdseks filtri pikkusega M , seega $L = M$. Sellise valiku korral saame algoritmi keerukuseks $O(2 \log_2(2M))$ aritmeetilist operatsiooni diskreedi kohta. Pikkade filtrite puhul annab see väiksema keerukuse kui LMS, mis mäletatavasti vajab $O(M)$ aritmeetilist operatsiooni diskreedi kohta.

Piiramata gradiendiga algoritm

Mõned autorid [13] on soovitanud algoritmi arvutusliku keerukuse edasiseks vähendamiseks jätta sagedusvalla kaalude rekursioonist (9.25) ära gradiendi piirangu $\mathbf{F} \begin{bmatrix} \mathbf{I} \\ \mathbf{0} \end{bmatrix} [\mathbf{I} \ \mathbf{0}] \mathbf{F}^H$ ning kasutada lihtsustatud algoritmi kujul

$$\begin{aligned} \mathbf{e}(t) &= \mathbf{d}(t) - \frac{1}{M + L} [\mathbf{0} \ \mathbf{I}] \mathbf{F}^H \Lambda \mathbf{f}(t) \\ \mathbf{f}(t + 1) &= \mathbf{f}(t) + \mu \mathbf{P}^{-1}(t) \Lambda^H \mathbf{F} \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} \mathbf{e}(t) \end{aligned} \tag{9.28}$$

Sellist modifikatsiooni kutsutaksegi piiramata gradiendiga sagedusvalla adaptiivseks algoritmiks. Sisuliselt asendab piiramata gradiendiga algoritm lineaarse korrelatsiooni kaalude rekursioonis tsirkulaarse korrelatsiooniga. Tuleb märkida, et selline operatsioon ei ole teoreetiliselt hästi põhjendatav, saadav algoritm võib töötada hästi mõningates ülesannetes, halvemini teistes ning seega nõuab ta enne kasutamist väga hoolikat testimist just vajalikus kontekstis.

Sagedusvalla algoritmid koonduvuse kiirendamiseks

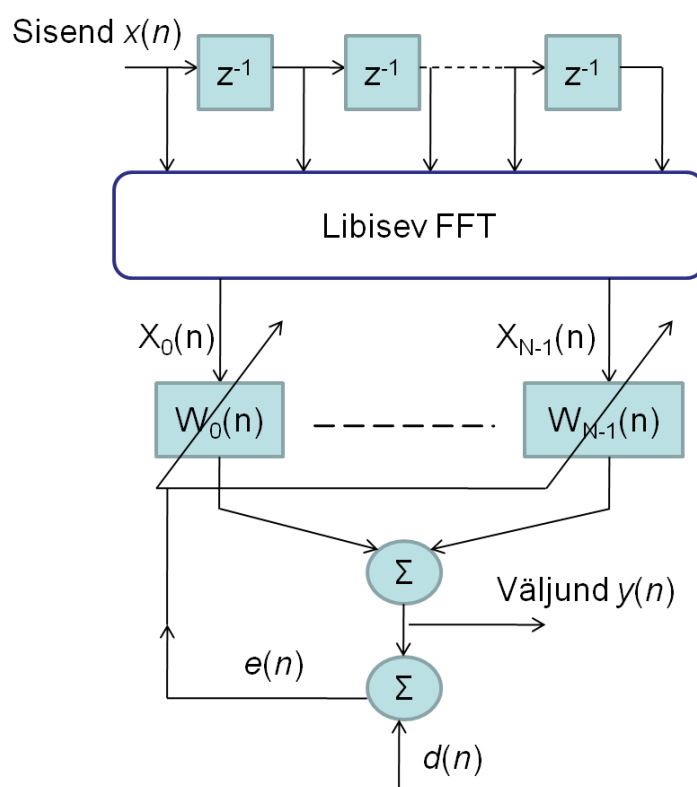
Üleminekut sagedusvalda on võimalik kasutada adaptiivse algoritmi koonduvuse kiirendamiseks, arvutusvõimsuse vähendamist taotlemata. Taoliste algoritmide positiivseks küljeks on ajalise viite vältimine. Sellised algoritmid lahutavad diskreetset Fourier' teisendust kasutades sisendsignaali kanalitesse ning rakendavad igas kanalis lihtsat ühe koefitsiendiga adaptiivset algoritmi. Kanalite

väljundsignaalid summeeritakse, saamaks adaptiivse filtri väljundsignaali. Olgu meil kasutusel N -punktiline libisev diskreetne Fourier' teisendus, mis produtseerib meile N spektri väärtust $X_i(n)$, $i = 0, \dots, N-1$ iga sisendsignaali diskreedi kohta. Paralleelsed adaptiivsed algoritmid rakendatakse seejärel spektrikomponentidele $X_i(n)$, tõlgendades neid kui sisendsignaale. Matemaatiliselt võime sellise adaptiivse filtri (pärast libisevat diskreetset Fourier' teisendust) võrrandid üles kirjutada kui

$$e(n) = d(n) - \sum_{i=0}^{N-1} w_i(n) X_i(n) \quad (9.29)$$

$$w_i(n+1) = w_i(n) + \mu_i X_i^*(n) e(n)$$

kusjuures μ_i on samm suurus i -ndas kanalis, mis valitakse pöördvõrdelisena võimsusega antud kanalis.



Joonis 60. Libiseval diskreetsel Fourier' teisendusel põhinev adaptiivne filter

Joonis 60 kujutab libiseval diskreetsel Fourier' teisendusel põhinevat adaptiivset filtrit graafiliselt. Näeme, et filtri väljundsignaali formeerimiseks liidetakse kokku kõikide adaptiivsete filterite väljundsignaalid ning veasignaali formeerimiseks lahutatakse tulemus soovitud signaalist. Selliselt saadud veasignaali adapteeritakse kõiki paralleelseid filterid. Selline konstruktsioon on võimalik, kuna iga diskreetse Fourier' teisenduse komplekssest eksponendist baasfunktsiooni st maatriksi \mathbf{F} iga rida, võime käsitleda, kui fikseeritud filtri impulsskarakteristikut

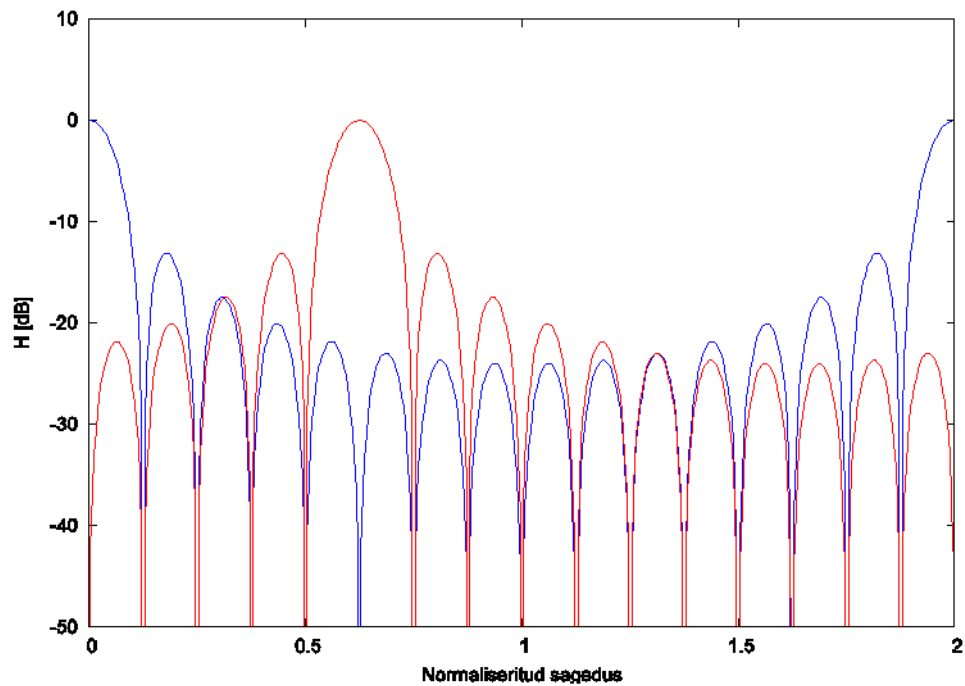
$$h(m) = \left[1, e^{-j2\pi m/N}, \dots, e^{-j2\pi(N-1)m/N} \right].$$

Ning libisevat diskreetset Fourier' teisendust

$$X(m) = \sum_{i=0}^{N-1} x(i) e^{-j2\pi im/N}, \quad m = 0, \dots, N-1 \quad (9.30)$$

võime vastavalt käsitleda kui paralleelselt ühendatud filtrite patareid impulsskarakteristikutega

$$h_m(n) = e^{-j2\pi mn/N}, \quad n = 0, \dots, N-1 \quad (9.31)$$



Joonis 61. Filtrite 0 ja 4 ülekandefunktsioonid

Filtrite ülekandefunktsiooni leiame kui z teisenduse tema impulsskarakteristikust, saades

$$H_m(z) = \sum_{i=0}^{N-1} e^{-j2\pi mi/N} z^{-i} \quad (9.32)$$

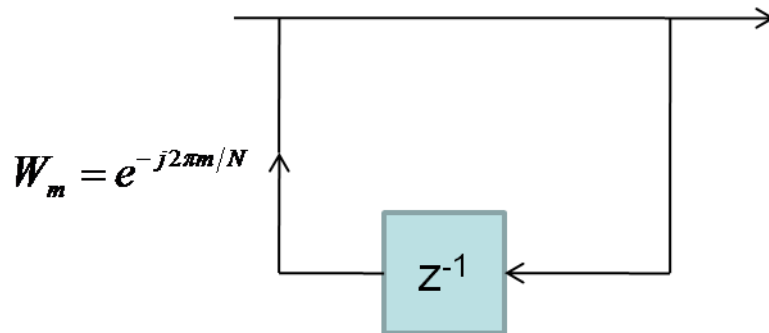
Edasi paneme tähele, et meie ülekandefunktsiooni valem kujutab endast N liikme geomeetrilist summat, mille võime arvutada kui

$$H_m(z) = \frac{1 - \left(e^{-j2\pi mi/N} z^{-1}\right)^N}{1 - e^{-j2\pi m/N} z^{-1}} \quad (9.33)$$

kuid kuna kompleksne eksponent 2π -kordsest argumendist on võrdne ühega, saame lõplikuks tulemuseks

$$H_m(z) = \frac{1 - z^{-N}}{1 - e^{-j2\pi m/N} z^{-1}} \quad (9.34)$$

Paneme tähele, et saadud ülekandefunktsiooni lugeja on kõikidele kanalitele ühine ning kanalid erinevad ainult nimetajas oleva koefitsiendi $e^{-j2\pi m/N}$ poolest.



Joonis 62. Patarei filter

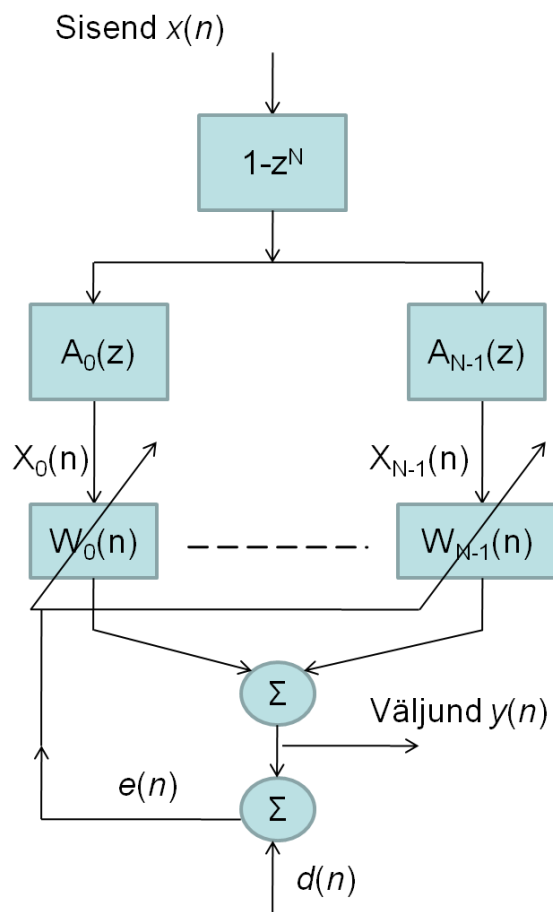
Joonis 62 kujutab seda osa kanali m filtrist, mis eristab antud kanalit kõikidest teistest. Tähistame selle

$$A_m(z) = \frac{1}{1 - e^{-j2\pi m/N} z^{-1}} \quad (9.35)$$

siis

$$H_m(z) = \frac{1 - z^{-N}}{1 - e^{-j2\pi m/N} z^{-1}} = (1 - z^{-N}) A_m(z). \quad (9.36)$$

Joonis 63 kujutab sellist rekursiivseid filtreid kasutavat adaptiivset filtrit graafiliselt. Paneme tähele, et filtreerimine läbi kõigile kanalitele ühise filtri $1 - z^{-N}$ toimub ühe korra, pärast mida jagatakse signaalid kanalitesse. Igas kanalis toimub seejärel filtreerimine individuaalse filtriga $A_m(z)$.



Joonis 63. Rekursiivseid filterid kasutatav adaptiivne algoritm

Lõppsõna

Kirjutada võiks veel paljudest huvitavatest asjadest, kuid kuna aeg selleks tööks on projekti peatse lõppemise tõttu otsas, siis piirdun sellega, mis hetkel kirjas. Loodan, et huvilistel on sellest ka kasu.

Indeks

ADPCM.....	7	maatriksi pööramise lemma	71
algoritm		meetod	
afiinse projektsiooni	81, 82, 83	kiireima languse	27, 28, 36, 79, 80
Newton-Raphsoni.....	78, 80, 91	vähimruutude	59, 60, 62, 66, 69, 81
sagedusvalla	83, 89, 91, 92, 94	operaator	
stohhastilise gradiendi	36	projektsiooni	63
ekvalaiser.....	6, 47, 48, 51, 52, 76	ortogonaalsuse printsiip	9
filter		piirang	
optimaalne	11	lineaarne	19
Wieneri	4, 8, 9, 16, 17, 32, 59	projektsioonimaatriks.....	63
Lagrange		pseudopöördmaatriks.....	66, 67, 68, 81
määramta kordajate meetod	22	pöördmudeli leidmine	6
lahutus singulaarväärtuste järgi	66	RLS.....	69, 70, 73, 75, 76, 77, 78, 81, 82, 83
lineaarne ennustamine.....	7	stabiilsus	9, 28, 29, 31, 78
LMF	55	süsteemi identifitseerimine	5, 43, 54, 56, 74, 75, 82, 92
LMS....	35, 36, 37, 39, 41, 42, 45, 53, 55, 76, 78	vea pind	14
lekkiv.....	55, 56	võrrandsüsteem	
märgi.....	55	Wiener-Hopfi	12
normaliseeritud	52, 53, 54, 75, 81	vähimruutude	65
plokk	58, 86, 88, 91	õppimisköver	33, 45, 46, 47, 75
maatriks			
Toeplitzi	14		

Kasutatud kirjandus

1. I. Arro, J. Derkatš. Digisignaalide töötlemine. TTÜ kirjastus, 2005.
2. J. M. Cioffi, T. Kailath. Windowed Fast Transversal Filters Adaptive Algorithms with Normalization. IEEE Trans. On Acoustics, Speech and Signal Processing, Vol. 33, No. 3, June 1985.
3. S. Gay, S. Tavathia. The Fast Affine Projection Algorithm. Proc. IEEE International Conference on Acoustics Speech and Signal Processing, 1995.
4. G.H. Golub, C.F. Van Loan. Matrix computations, third edition. The John Hopkins University Press, 1996.
5. S. Haykin. Adaptive filter theory, fourth edition. Prentice Hall, 2002.
6. ITU-T Recommendation G.726. 40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation.
7. C.L. Lawson, R.J. Hanson. Solving Least Squares Problems. Prentice Hall, 1974.
8. S.L. Marple, Jr. Digital spectral analysis with applications. Prentice Hall, 1987.
9. A.V. Oppenheim, R.W. Schaffer. Discrete-Time Signal Processing. Prentice-Hall 1989.
10. A. H. Sayed. Fundamentals of adaptive filtering. J. Wiley and sons 2003.
11. D.T. Slock, T. Kailath. Numerically stable fast transversal filters for recursive least squares adaptive filtering. IEEE Trans. On Signal Processing, Vol. 39, No. 1, January 1991.
12. P. Stoica, R. Moses. Spectral Analysis of Signals. Prentice Hall, 2005.
13. J. Shynk. Frequency-Domain and Multirate Adaptive Filtering. IEEE Signal Processing Magazine, Jan. 1992.
14. E. Tamme, L. Võhandu, L. Luht. Arvutusmeetodid. Valgus, 1986.
15. I. Tammeraid. Lineaaralgebra rakendused. TTÜ kirjastus, 1999.
16. S. Theodoridis, M. Bellanger. Adaptive Filters and Acoustic Echo Control. IEEE Signal Processing Magazine, July, 1999.
17. B. Widrow, S. D. Stearns „Adaptive signal processing,“ Prentice-Hall, 1985.
18. B. Widrow, M.E. Hoff. Adaptive switching circuits. IRE WESCON Conv. Rec., Pt. 4, pp. 96 – 104.
19. N. Yousef, A. Sayed. A Unified Approach to the Steady State and Tracking Analysis of Adaptive Filters. IEEE Trans. On Signal Processing, Feb. 2001.
20. K. Zhao, F. Ling, H. Lev-Ari, J.G. Proakis. Sliding Window Order – Recursive Least – Squares Algorithms. IEEE Transactions on Signal Processing, Vol. 4, No. 8, August 1994.