Helsinki University of Technology
Laboratory of Telecommunications Technology

S-38.211 Signaalinkäsittely tietoliikenteessä I
Signal Processing in Communications (2 ov)

Syksy 1998

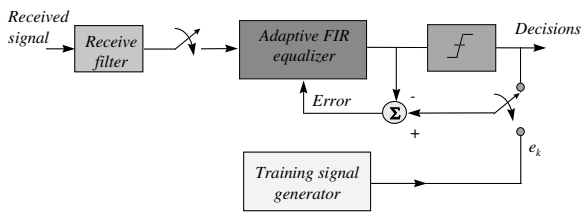
6. Luento: Adaptiiviset korjaimet I

prof. Timo Laakso
Vastaanotto torstaisin klo 10-11
Huone G210, puh. 451 2473
Sähköposti: timo.laakso@hut.fi

Adaptiivinen FIR-korjain (LM 11)

- Viimeksi tarkasteltiin ideaalisen korjaimen *z*-siirtofunktiota. Z-siirtofunktion kompleksisuutta ei rajoitettu, ja se oli yleensä paitsi rekursiivista muotoa myös ääretön asteluvultaan.
- Seuraavassa rajoitutaan *äärellisen pituiseen FIR-korjaimeen*. Ensin johdetaan matriisinotaation avulla *ideaalinen MMSE-ratkaisu*. Tämän pohjalta kehitetään *gradienttityyppinen iteratiivinen ratkaisumenetelmä* matriisiyhtälölle joka yksinkertaistuu aikanaan käyttökelpoiseksi *LMS (least mean square)-adaptointi-algoritmiksi*.
- LMS-algoritmi on pohjana useimmille käytännön lineaarisille ja takaisinkytketyille (DFE-)korjaimille.

Adaptiivisen korjaimen rakenne



- Adaptiivisen *lineaarisen* korjaimen perusrakenne on **Kuvassa 11-1**.

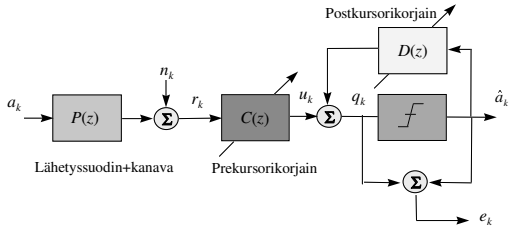
...Adaptiivisen lin. korjaimen rakenne

- Alkupään vastaanottosuodatin on yleensä *alipäästösuodatin*, joka poistaa kohinaa
- Adaptiivinen (lineaarinen) korjain toteutetaan yleensä *FIR-rakenteella*
- Korjaimen kertoimia ohjataan *virhesignaalilla.*

Virhesignaali muodostetaan

- *Alustusvaiheessa* (training mode) lähettämällä tunnettu testisignaali
- *Seurantavaiheessa* (tracking mode) käyttämällä hyväksi menneitä päätöksiä. Virheelliset päätökset (alle 10%) eivät juuri haittaa.

Adaptiivisen DFE-korjaimen rakenne



- Päätöstakaisinkytkennnän sisältävän kantataajuisen DFE-korjaimen diskreettiaikainen mallin perusrakenne on **Kuvassa 11-2**.

...Adaptiivisen DFE-korjaimen rakenne

- Adaptointi perustuu virhesignaaliin missä on *alustusvaiheessa* tunnettu symbolisekvenssi ja *seurantavaiheessa* vastaanottimen päätössekvenssi.
- *Prekursorin N*-tappinen siirtofunktio on muotoa (Huom! Kanava-lähetyssuodin on nyt $P(z)$ ja korjain $C(z)$)

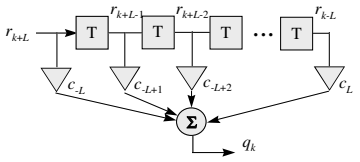
$$C(z) = \sum_{m=-(N-1)}^{0} c_m z^{-m} \qquad (11.3)$$

(käytännössä kausaloitava N viiveellä). *Postkursori* on

$$D(z) = \sum_{m=1}^{M} d_m z^{-m} \qquad (11.4)$$

- Tämä toteutetaan M-tappisella FIR-rakenteella.

FIR-korjaimen MSE-ratkaisu (LM 11.1.2)



- Johdetaan N-tappisen FIR korjaimen (11.2) MSE. Määritellään ensin kerroinvektori

$$\mathbf{c} = \begin{bmatrix} c_{-L} & \cdots & c_0 & \cdots & c_L \end{bmatrix}^{\mathrm{T}} \qquad (11.5)$$

- missä ovat FIR-suotimen kertoimet, $N = 2L+1$ ja N oletetaan parittomaksi.

... FIR-korjaimen MSE-ratkaisu

- Vastaavat suodatettavat korjaimen input-näytteet hetkellä k kootaan samanmittaiseksi vektoriksi:

$$\mathbf{r}_k = \begin{bmatrix} r_{k+L} & \cdots & r_k & \cdots & r_{k-L} \end{bmatrix}^{\mathrm{T}} \qquad (11.6)$$

- Virhesignaali saadaan muotoon

$$e_k = a_k - q_k = a_k - \mathbf{c}^{\mathrm{T}}\mathbf{r}_k \qquad (11.7)$$

- Yleisesti kaikki suureet ovat *kompleksiarvoisia*. Oletetaan lisäksi että kaikki signaalit ovat *nollakeskiarvoisia* ja (laajassa mielessä) *stationäärisiä* eli niiden stokastiset ominaisuudet eivät muutu ajan funktiona.

... FIR-korjaimen MSE-ratkaisu

- Neliövirhe voidaan työstää muotoon (**Harjoituksissa!**)

$$\mathrm{E}\left[|e_k|^2\right] = \mathrm{E}\left[|a_k|^2\right] - 2\,\mathrm{Re}\left\{\mathbf{c}^{\mathrm{H}}\mathrm{E}\left[a_k\mathbf{r}_k^*\right]\right\} + \mathbf{c}^{\mathrm{H}}\mathrm{E}\left[\mathbf{r}_k\mathbf{r}_k^{\mathrm{H}}\right]\mathbf{c}$$
$$= \mathrm{E}\left[|a_k|^2\right] - 2\,\mathrm{Re}\left\{\mathbf{c}^{\mathrm{H}}\mathbf{a}\right\} + \mathbf{c}^{\mathrm{H}}\boldsymbol{\Phi}\mathbf{c} \qquad (11.8)$$

- missä yläindeksi 'H' merkitsee *hermitointia* eli transponointia ja konjugointia. Lisäksi määritellään *ristikorrelaatiovektori*

$$\mathbf{a} = \mathrm{E}\left[a_k\mathbf{r}_k^*\right] \qquad (11.9)$$

- ja *autokorrelaatiomatriisi*

$$\boldsymbol{\Phi} = \mathrm{E}\left[\mathbf{r}_k\mathbf{r}_k^{\mathrm{H}}\right] = \begin{bmatrix} \Phi_0 & \Phi_{-1} & \cdots & \Phi_{-(N-1)} \\ \Phi_1 & \Phi_0 & & \\ \vdots & & \ddots & \vdots \\ \Phi_{N-1} & & \cdots & \Phi_0 \end{bmatrix} \qquad (11.10)$$

Autokorrelaatiomatriisi

- **Autokorrelaatiomatriisilla** on tärkeitä ominaisuuksia:
- 1) on *Hermiten (hermiittinen) matriisi*, eli $\boldsymbol{\Phi}^{\mathrm{H}} = \boldsymbol{\Phi}^{\mathrm{T}*} = \boldsymbol{\Phi}$ (jos $\boldsymbol{\Phi}$ on reaalinen, se on symmetrinen).
- 2) on *Töplitz-matriisi* eli diagonaalin suuntaisilla alidiagonaaleilla on sama elementti. *Autokorrelaatiofunktion* arvo riippuu siis vain korreloitavien viive-erosta: $\boldsymbol{\Phi}_j = \mathrm{E}[r_{k+j}r_k^*]$
- 3) on *positiivi-semidefiniitti* eli mille tahansa vektorille $\mathbf{x}$ $\mathbf{x}^{\mathrm{H}}\boldsymbol{\Phi}\mathbf{x}$ on positiivinen ja on siten tulkittavissa *neliölliseksi pituusmitaksi*.

Optimaalinen MMSE-ratkaisu

- Tarkastellaan kaavan (11.8) neliövirhelauseketta kun FIR-suotimen pituus on 1 ja signaalit ovat reaaliarvoisia. Saadaan

$$\mathrm{E}\left[|e_k|^2\right] = \mathrm{E}\left[|a_k|^2\right] - 2ac + \Phi_0 c^2$$
$$a = \mathrm{E}\left[a_k r_k\right], \quad \Phi_0 = \mathrm{E}\left[r_k^2\right] \qquad (11.14)$$

- Tämä voidaan helposti täydentää muotoon

$$\mathrm{E}\left[|e_k|^2\right] = \mathrm{E}\left[|a_k|^2\right] - a^2/\Phi_0 + \Phi_0\left(a^2/\Phi_0^2 - 2ac/\Phi_0 + c^2\right)$$
$$= \mathrm{E}\left[|a_k|^2\right] - a^2/\Phi_0 + \Phi_0\left(c - a/\Phi_0\right)^2$$

josta nähdään että MSE:n minimoiva ratkaisu c:lle on …

...Optimaalinen MMSE-ratkaisu

- Tämä yksinkertainen tulos on voidaan yleistää kompleksiseen vektoritapaukseen:

LM Harj. 11-4 Johda kertomalla auki, että

$$\mathrm{E}\left[|e_k|^2\right] = \mathrm{E}\left[|a_k|^2\right] - \mathbf{a}^{\mathrm{H}}\boldsymbol{\Phi}^{-1}\mathbf{a} + (\boldsymbol{\Phi}^{-1}\mathbf{a} - \mathbf{c})^{\mathrm{H}}\boldsymbol{\Phi}(\boldsymbol{\Phi}^{-1}\mathbf{a} - \mathbf{c}) \qquad (11.15)$$

on yhtäpitävä (11.8:n kanssa).

- Koska $\boldsymbol{\Phi}$ on positiivi-semidefiniitti, kaavan (11.15) viimeinen termi on aina ei-negatiivinen eli pienimmillään nolla. Optimiratkaisu saadaan siis valitsemalla

$$\mathbf{x} = \boldsymbol{\Phi}^{-1}\mathbf{a} - \mathbf{c} = \mathbf{0} \quad \Leftrightarrow \quad \mathbf{c}_{\mathrm{opt}} = \boldsymbol{\Phi}^{-1}\mathbf{a} \qquad (11.16)$$

- Ratkaisu (jos se on olemassa) on aina *yksikäsitteinen*.

...Optimaalinen MMSE-ratkaisu

- Edelläsaatu ratkaisu (11.16) minimoi lausekkeen (11.15) MSE:n, sillä muut termit eivät riipu kerroinvektorista **c**. Minimi-MSE:ksi saadaan

$$\xi_{\min} = \mathrm{E}\left[|a_k|^2\right] - \mathbf{a}^{\mathrm{H}}\Phi^{-1}\mathbf{a} = \mathrm{E}\left[|a_k|^2\right] - \mathbf{a}^{\mathrm{H}}\mathbf{c}_{opt} \qquad (11.17)$$

- MSE:n yleinen kaava saadaan siis muotoon

$$\mathrm{E}\left[|e_k|^2\right] = \xi_{\min} + (\Phi^{-1}\mathbf{a} - \mathbf{c})^{\mathrm{H}}\Phi(\Phi^{-1}\mathbf{a} - \mathbf{c}) \qquad (11.18)$$

- MSE riippuu suoraan siitä, kuinka lähellä **c** on optimia. MSE:n kuvaaja on *(hyper)paraboloidi N*-ulotteisessa avaruudessa.
- Optimiratkaisu (11.16) voidaan johtaa myös *derivoimalla* MSE:n lauseke vektorin **c** suhteen ja asettamalla derivaatta nollaksi (ks. kirjan **Harj. 11-5**). Tulos on tietysti sama.

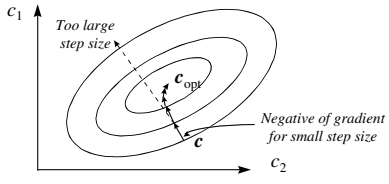
MSE-gradienttialgoritmi (LM 11.1.3)

- Kaavan (11.16) optimikerroinvektori löydetään *matriisinkäännöllä* eli ratkaisemalla lineaarinen yhtälöryhmä joka riippuu autokorrelaatio*matriisin* Φ termeistä (autokorrelaatio*funktion* Φj arvoista) ja ristikorrelaatio-vektorista **a**. Jos kerroinvektoria joudutaan päivittämään usein (niinkuin oletetaan), se on työläs tapa, ja lisäksi siinä voi tulla numeerisia (laskentatarkkuus-)ongelmia.
- Lisäksi autokorrelaatio- ja ristikorrelaatiotermit täytyy tuntea (tai estimoida).

...MSE-gradienttialgoritmi

- Suoran ratkaisun sijasta etsimme *iteratiivista* algoritmia, joka vähitellen lähestyy optimiratkaisua. Ratkaisu perustuu siihen perustulokseen että optimikerroinvektori on *yksikäsitteinen ja* virhepinta (MSE:n kuvaaja kerroinvektoriavaruudessa) on *unimodaalinen*, eli sillä on vain yksi globaali minimi (ei paikallisia ääriarvoja!).
- Tällöin voidaan MSE-käyrää pitkin edetä aina negatiivisen gradientin suuntaan eli *alamäkeen* optimiratkaisua kohti.

...MSE-gradienttialgoritmi



- **Kuva 11-3** havainnollistaa gradienttialgoritmin periaatteellista toimintaa (kaksiulotteinen tapaus).

...MSE-gradienttialgoritmi

- *MSE-gradienttialgoritmi (MSEG)* voidaan esittää muodossa

$$\mathbf{c}_{j+1} = \mathbf{c}_j - \frac{\beta}{2}\nabla_{\mathbf{c}_j}\mathrm{E}\left[|e_k|^2\right] \qquad (11.27)$$

- missä $\mathbf{c}_{j+1}$ ja $\mathbf{c}_j$ ovat uusi ja vanha kerroinvektori, β on askelparametri (1/2 on mukana jotta seuraavat kaavat sievenevät) ja nablalauseke on MSE-virhepinnan gradientti. Derivoimalla MSE:n lauseke saadaan (ks. kirjan **Harj. 11-5**)

$$\nabla_{\mathbf{c}_j}\mathrm{E}\left[|e_k|^2\right] = 2\Phi\mathbf{c} - 2\mathbf{a} \qquad (11.25)$$

- jolloin algoritmi (11.27) saadaan muotoon

$$\mathbf{c}_{j+1} = \mathbf{c}_j + \beta(\mathbf{a} - \Phi\mathbf{c}_j) = (\mathbf{I} - \beta\Phi)\mathbf{c}_j + \beta\mathbf{a} \qquad (11.28)$$

...MSE-gradienttialgoritmi

- **Esimerkki 11-2**: tutkitaan taas *reaalista skalaaritapausta* (**Esim. 11-1**), jolloin MSEG-algoritmi saadaan muotoon

$$c_{j+1} = (1 - \beta\Phi_0)c_j + \beta a \qquad (11.28)$$

- Vähennetäänpä kummaltakin puolelta optimiratkaisu c_{opt} ja käytetään kaavaa $a=\Phi_0 c_{\mathrm{opt}}$:

$$\begin{aligned} c_{j+1} - c_{opt} &= (1 - \beta\Phi_0)c_j - c_{opt} + \beta\Phi_0 c_{opt} \\ &= (1 - \beta\Phi_0)(c_j - c_{opt}) \end{aligned} \qquad (11.30a)$$

Kun oletetaan kertoimen alkuarvoksi , voidaan (*j*+1*)*:s ratkaisu lausua myös muodossa

$$c_{j+1} - c_{opt} = (1 - \beta\Phi_0)^j(c_0 - c_{opt}) \qquad (11.30b)$$

...MSE-gradienttialgoritmi

Tästä kaavasta nähdään, että

- Ratkaisu suppenee *aina* kohti optimia (alkuarvosta riippumatta) jos

$$|1 - \beta\Phi_0| < 1$$

- Tästä seuraa *askelparametrille* vaatimus

$$-1 < \beta\Phi_0 - 1 < 1 \quad \Leftrightarrow \quad 0 < \beta\Phi_0 < 2 \quad \Leftrightarrow \quad 0 < \beta < \frac{2}{\Phi_0}$$

- Nopeimmin algoritmi suppenee jos $\beta=1/\Phi_0$: tällöin tarvitaan vain yksi iteraatio!

MSEG-algoritmin suppeneminen

- Edellisen esimerkin tulos voidaan yleistää kompleksiarvoiseen vektoritapaukseen. Esitetään kaava (11.28) muodossa

$$\begin{aligned} \mathbf{c}_{j+1} - \mathbf{c}_{opt} &= (\mathbf{I} - \beta\Phi)\mathbf{c}_j - \mathbf{c}_{opt} + \beta\Phi\mathbf{c}_{opt} \\ &= (\mathbf{I} - \beta\Phi)(\mathbf{c}_j - \mathbf{c}_{opt}) \end{aligned}$$

- Merkitsemällä $\mathbf{q}_j=\mathbf{c}_j-\mathbf{c}_{\mathrm{opt}}$ ja käyttämällä alkuarvoa $\mathbf{q}_0$ saadaan ratkaisu muotoon

$$\mathbf{q}_{j+1} = (\mathbf{I} - \beta\Phi)^j\mathbf{q}_0 \qquad (11.32)$$

- Mutta milloin algoritmi suppenee? Nyt sen selvittäminen on hieman työläämpää kuin skalaaritapauksessa ja matriisin ominaisarvoja ja -vektoreita tarvitaan.

...MSEG-algoritmin suppeneminen

- Kaavan (11.32) kerroinmatriisi voidaan esittää *modaalihajotelmana*

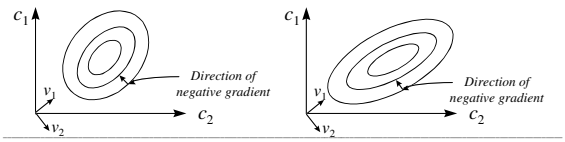
$$(\mathbf{I}-\beta\Phi)^j=\sum_{i=1}^{N}(1-\beta\lambda_i)^j\mathbf{v}_i\mathbf{v}_i^{\mathrm{H}} \qquad (11.38)$$

- missä λ_i on matriisin Φ *ominaisarvo* (aina reaalinen ja ei-neg.!) ja $\mathbf{v}_i$ vastaava *ominaisvektori* ($\mathbf{v}_i^{\mathrm{H}}\mathbf{v}_i=1$). Termi $(1-\beta\lambda_i)^j$ on lausekkeen (11.38) *i*:s *suppenemismoodi*.
- Jotta jokainen moodi suppenisi, on vaadittava että *suurinta* ominaisarvoa vastaava moodi suppenee, mistä seuraa *askelparametrin suppenemisehto*

$$0<\beta<\frac{2}{\lambda_{\max}} \qquad (11.39)$$

...MSEG-algoritmin suppeneminen

- Voidaan lisäksi osoittaa, että suppeneminen on hidasta kun ns. ominaisarvohaje $\lambda_{\max}$ / $\lambda_{\min}$ (suurimman ja pienimmän ominaisarvon suhde) on suuri.
- Käytännössä suuri ominaisarvohaje ilmenee siten, että MSE-pinnan (hyper)paraboloidin poikkipinta on litistynyt (hyper)ellipsi. Tällöin *gradienttisuunta kääntyy voimakkaasti* edettäessä kohti maljan pohjaa, jolloin on pakko käyttää pienehköä askelparametria ettei ajeta ohi. **Kuva 11-6**:



Ominaisarvot ja tehospektri

- Vieläkin havainnollisempi on spektritulkinta. Voidaan osoittaa, että *signaalin tehospektri asettaa rajat ominaisarvohajeelle:*

$$\min_{\omega}\left[S(e^{j\omega})\right]<\lambda<\max_{\omega}\left[S(e^{j\omega})\right] \qquad (11.45)$$

ja tehospektrihän on *autokorrelaatiofunktion Fourier-muunnos*:

$$S(e^{j\omega})=\sum_{k=-\infty}^{\infty}\Phi_k e^{-j\omega k} \qquad (11.46)$$

- Ominaisarvot riippuvat matriisin dimensioista. Voidaan osoittaa, että ominaisarvojen minimi ja maksimi lähestyvät (11.45):n raja-arvoja kun matriisin kokoa (eli adaptoitavan *suotimen pituutta*) kasvatetaan.

...Ominaisarvot ja tehospektri

- Suppenemisen kannalta on siten parasta, että korjaimen input-signaalin *spektri on mahdollisimman tasainen* (lähellä valkoista kohinaa). Jos spektrissä on yksikin nollakohta, suppeneminen hidastuu periaatteessa nollaan kun suotimen pituutta kasvatetaan.
- Käytännön pohjakohinan takia tämä on lähinnä teoreettinen ongelma. Silti *pitkien suotimien adaptointi on hidasta* varsinkin kun signaalin spektri on vaihteleva.

Yhteenveto 6. luennosta

Tällä luennolla käsiteltiin FIR-korjaimen adaptointia. Tutustuttiin mm. seuraaviin käsitteisiin

- MSE-virheen lauseke *matriisimuodossa* *N*-pituiselle FIR-suotimelle
- MSE-gradienttialgoritmi (MSEG)
- MSEG-algoritmin suppeneminen
- Askelparametrin määrittäminen
- Ominaisarvohaje, tehospektri ja suppeneminen