

Adaptív digitális szűrés I.

DR. VARGA IMRE

BME Híradástechnikai Elektronika Intézet

ÖSSZEFOGLALÁS

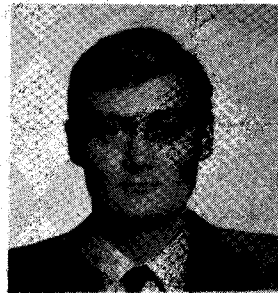
A cikk két részben áttekintést nyújt az adaptív szűrők struktúráiról, algoritmusairól és bemutatja a jellegzetes problémákat. Az első rész az optimális lineáris szűréssel foglalkozik, részletesen tárgyalja azokat a FIR szűrő algoritmusokat, melyek a Wiener-Hopf egyenletből és a Kalman-szűrőből vezethetők le.

1. Bevezetés

Jelen két részes cikkben szűrő alatt olyan jelfeldolgozó eszközt értünk, amely egy stochasztikus folyamatból egy másikat valamilyen hibakritérium értelemben lineáris szűréssel állít elő. Alapvetően háromféle információfeldolgozási folyamatról beszélhetünk: szűrésről van szó, ha a t időpillanatban és előtte, simításról, ha a t időpillanat előtt és után mért adatokból következtünk a mennyiség t időpontbeli értékére, míg predikció esetén a t időpillanatban és előtte mért adatokból jósoljuk a mennyiség értékét egy későbbi, $t_1 > t$ időpontban.

Az optimális lineáris szűrés feladata annak a lineáris szűrőnek a meghatározásában áll, mely egy adott stochasztikus folyamatból egy másik adott lineáris szűréssel, adott hibakritérium értelemben minimális hibával elő tud állítani. A hibakritérium szokásosan a négyzetes átlaghiba minimalizálása és az így kapott optimális lineáris nemrekurzív becslő a Wiener-szűrő stacionárius folyamatok esetén, illetve az optimális lineáris rekurzív becslő a Kalman-szűrő, mely nemstacionárius folyamatokra is alkalmazható. A gyakorlatban az optimális szűrő számításához szükséges a priori információ a feldolgozandó jelekről nem áll rendelkezésre, s ezért van szükség adaptív szűrőkre. Az adaptív szűrők "öntervezők" abban az értelemben, hogy a bemeneti stochasztikus folyamatok statisztikájának ismerete nélkül, egy rekurzív algoritmus - bizonyos kezdeti feltételekből kiindulva - iteratív módon gondoskodik az optimális közelálló szűrőkarakterisztika illetve szűrőegység kialakításáról. Az algoritmus rekurzív jellege miatt a kimeneti jel függ a bemeneti jeltől, vagyis az adaptív szűrő általában nemlineáris eszköz.

Az adaptív szűrők elvileg két független részből, algoritmusból és struktúrából tevődnek össze. A cikkben ismertetjük az adaptív FIR (véges impulzusválaszú) szűrőkre vonatkozó algoritmusokat, majd az adaptív IIR (végtelen impulzusválaszú) szűrők jellegzetes problémáit tekintjük át.



Dr. VARGA IMRE

A BME Villamosmérnöki Karán végzett 1982-ben. 1982-1984. között a BME Elméleti Villamosságtan Tanszéken volt tudományos továbbképzési ösztöndíjas, a hálózatelmélet témakörén belül lineáris hálózatok érzékenység- és toleranciaproblémáival foglalkozott. 1984-ben műszer és irányítástechnikai szakmérnöki és egyetemi doktori oklevelet szerzett. 1984. óta a BME Híradástechnikai Elektronika Intézetben dolgozik. Jelenlegi fő érdeklődési területe a digitális jelfeldolgozás, adaptív szűrés.

2. Stochasztikus approximáció

A stochasztikus approximáció módszerei jelentik történetileg az első lépést adaptív algoritmus kialakítására. Ezek az eljárások rekurzívak, a legkisebb négyzetek módszerénél egyszerűbbek, konvergenciájuk ugyanakkor kevésbé megbízható. A Robbins-Monro algoritmus [8] tárgyalásához tekintsük az

$$f(h) = f_0 \quad (2-1)$$

egyenletet, melyben f_0 ismert konstans és a feladat a h paraméter meghatározása. Abban az esetben, ha $f(h)$ pontosan mérhető, (2-1) megoldható pl. a jól ismert gradiensmódszerrel:

$$h_{k+1} = h_k + \mu_k [f(h_k) - f_0] \quad (2-2)$$

Az iteráció bizonyos feltételeket kielégítő μ_k számsorozat esetén konvergens.

Abban az esetben, ha $f(h)$ helyett annak csak zajjal terhelt

$$f_n(h, n) = f(h) + n \quad (2-3)$$

értéke mérhető, ahol n zérus középértékű, véges varianciájú zajt jelent és

$$E\{f_n(h, n)\} = f(h) + n \quad (2-4)$$

(E a várható érték képzést jelöli), akkor a gradiensmódszer alkalmazása a

$$\hat{h}_{k+1} = \hat{h}_k + \mu_k [f_n(\hat{h}_k, n_k) - f_0] \quad (2-5)$$

Robbins-Monro iterációra vezet, mely a (2-4) egyenlet h_0 megoldásához konvergál legkisebb átlagos négyzetes hiba értelemben, ha a (2-4) egyenletnek egy gyöke van és

$$\lim_{k \rightarrow \infty} \mu_k = 0, \quad \sum_{k=1}^{\infty} \mu_k = \infty, \quad \sum_{k=1}^{\infty} \mu_k^2 < \infty. \quad (2-6)$$

Példa a fenti feltételeknek eleget tevő μ_k számsorozatra: $\mu_k = \alpha/k$ vagy $\mu_k = \alpha/(\beta+k)$. Az algoritmus neve abból származik, hogy a (2-5) rekurzióban a pontos gradiens "stochasztikusan" közelítő értéke szerepel.

Ehhez hasonló a Kiefer-Wolfovitz algoritmus, mely az $f(h)$ függvény szélsőérték helyét, tehát a

$$\frac{df(h)}{dh} = 0 \quad (2-7)$$

egyenlet megoldását közelíti a

$$\hat{h}_{k+1} = \hat{h}_k + \mu_k \frac{d}{dh} f(\hat{h}_k, n_k) \quad (2-8)$$

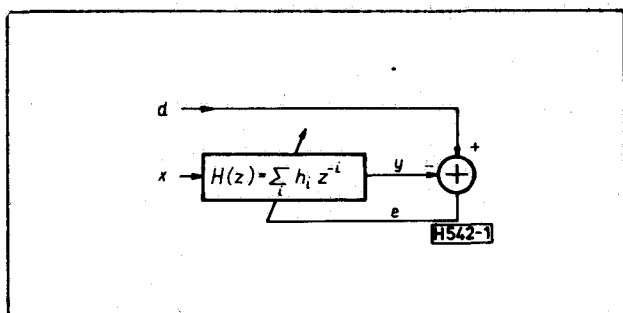
rekurzióval. A stochasztikus approximáció módszereinek konvergenciájával Blum (1953), Dvoretzky (1956) és Derman és Sacks (1959) foglalkozott részletesen.

A stochasztikus approximáció (2-5) módszere

$\lim_{k \rightarrow \infty} \mu_k = 0$ miatt végtelen memóriájú és emiatt stacionárius jeleknél alkalmazható.

3. Wiener-Hopf egyenlet és LMS (least mean squares) algoritmus

Az 1. ábrán feltüntetett rendszerben az x bemeneti jel lineáris szűrésével olyan y kimeneti jelet kívánunk előállítani, amely legkisebb átlagos négyzetes hiba



1. ábra Adaptív szűrő

értelemben közelíti a d (desired response) bemeneti jelet. Feltéve, hogy x és d valós értékű stacionárius folyamat, akkor

$$\begin{aligned} e &= d - y \\ y &= h^T x \end{aligned} \quad (3-1)$$

$$J(h) = E\{e^2\} \rightarrow \min$$

ahol az x jel egymás utáni értékeit x -ba foglaltuk össze és ezzel a helyettesítéssel

$$J(h) = \sigma_{2d} - 2h^T p + h^T R h \rightarrow \min \quad (3-2)$$

adódik, ahol

$$R = E\{xx^T\}, \quad p = E\{dx\} \quad (3-3)$$

a bemeneti jel autokorrelációs mátrixa és a két bemeneti jel keresztkorrelációs vektora. A (3-2) hibafüggvény minimalizálása ekvivalens az

$$Rh = p \quad (3-4)$$

lineáris egyenletrendszer (Wiener-Hopf egyenlet vagy normálegyenlet) megoldásával. Az optimális

$$h_{opt} = R^{-1} p \quad (3-5)$$

szűrőegyütthatóvektor általában az időinvariáns és nemkauzális Wiener-szűrőt írja le, azonban ha hosszát speciális megkötésként véges N -re választjuk úgy, hogy $H(z)$ FIR szűrőt jelentsen:

$$h(i) = \begin{cases} 0, & i < 0 \\ \text{véges}, & 0 \leq i \leq N \\ 0, & i > N \end{cases} \quad (3-6)$$

akkor h_{opt} az optimális FIR szűrő együtthatóit jelöli. A feltételezett stacionárius esetben tehát a Wiener-szűrő időinvariáns, h_{opt} időben nem változik.

Az optimális lineáris Wiener-szűrő lényeges tulajdonsága az ún. ortogonalitás elve. Átrendezve ugyanis a Wiener-Hopf egyenletet (3-4),

$$E\{x_k(d_k - h^T_{opt} x_k)\} = E\{x_k e_{optk}\} = 0$$

adódik, tehát a FIR szűrő jelvektorának elemei és az optimális helyettséghez tartozó hibák ortogonálisak. Hasonlóképp, felírva az előbbi egyenlet h^T_{opt} vektorral beszorzott alakját,

$$h^T_{opt} E\{x_k e_{optk}\} = 0 \quad (3-7)$$

adódik, vagyis

$$E\{y_k e_{optk}\} = 0 \quad (3-8)$$

Ez azt jelenti, hogy a szűrő kimenete és a hiba is ortogonális, vagyis e normálisa y -nak. Ezért a Wiener-Hopf egyenletet normálegyenletnek is nevezik. A minimális átlagos négyzetes hiba és a bemenet-hiba közötti ortogonalitás tehát azonos feltételeket jelent.

A Wiener-Hopf egyenlet megoldásának problémái három szinten jelentkeznek. Egyrészt ha a priori ismert is a bemeneti jelek statisztikája, vagyis az R és p mennyiségek, akkor is $O(N^3)$ számításigényt jelent a

(3-5) direkt megoldás. Ez $O(N^2)$ -re redukálható, kihasználva, hogy stacionárius esetben \mathbf{R} szimmetrikus Töplitz mátrix. Másrészt, a gyakorlatban \mathbf{R} és \mathbf{p} ismeretlen. Bár lehetőség van az adatok ablakolásával, blokkonként becsülni a mennyiségeket és megoldani a Wiener-Hopf egyenletet, ennek egy rekurzív algoritmussal szemben nincs előnye. Harmadrészt, a gyakorlatban a jeleknek nem csak a statisztikája ismeretlen, hanem fel kell készülni nemstacionárius helyzet követésére is. Mindhárom szempontból az $O(N)$ számításigényű, rendkívül egyszerű rekurzív LMS algoritmus kedvezőbb, mely a gradiensmódszer "stochasztikus változata".

Miután

$$\text{grad } J(\mathbf{h}) = -2\mathbf{p} + 2\mathbf{R}\mathbf{h} \quad (3-9)$$

így a gradiensmódszer iterációja

$$\mathbf{h}_{k+1} = \mathbf{h}_k + \mu(\mathbf{p} - \mathbf{R}\mathbf{h}_k), \quad (3-10)$$

ahol μ a lépésmagyságot jelöli.

Ha a jelek statisztikája ismeretlen, akkor a gradienszt illetve az \mathbf{R} és \mathbf{p} mennyiségeket becsülni kell. Az LMS algoritmus (Windrow és Hoff, 1960) [5], [6], [10], [9] az alábbi közelítést teszi:

$$\mathbf{R} = E\{\mathbf{x}\mathbf{x}^T\} \rightarrow \hat{\mathbf{R}}_k = \mathbf{x}_k\mathbf{x}_k^T \quad (3-11)$$

$$\mathbf{p} = E\{d\mathbf{x}\} \rightarrow \hat{\mathbf{p}}_k = d_k\mathbf{x}_k$$

Így tehát az LMS rekurzió

$$\mathbf{e}_k = d_k - \mathbf{h}_k^T \mathbf{x}_k \quad (3-12)$$

$$\mathbf{h}_{k+1} = \mathbf{h}_k + \mu \mathbf{e}_k \mathbf{x}_k \quad (3-13)$$

alakú. Két adat érkezése közötti időben tehát egyrészt szűrést (3-12), másrészt adaptációt (3-13) kell végezni. Az LMS algoritmust stochasztikus gradiensmódszernek is nevezik, mert a gradiensnek pillanatnyi, "stochasztikus" közelítését alkalmazza, annak ellenére, hogy éppen μ közelítés miatt a gradiens helyén véletlen számok állnak. Inicializálásként általában $\mathbf{h}_0 = 0$ megfelelő.

Az alábbiakban összefoglaljuk a LMS algoritmus főbb tulajdonságait [5], [6], [9], [10].

1. Az LMS algoritmus műveletigénye $O(N)$, pontosabban a szorzások iterációnkénti száma $2N$. Legfőbb előnye éppen az egyszerűsége.

2. Stabilitás. A $J(\mathbf{h})$ hibafelület az N dimenziós térben stacionárius esetben egy rögzített alakú és helyzetű paraboloid, melynek globális minimuma keresett. Ha a jelek statisztikája ismert, akkor a paraboloid is ismert és minimumát például a gradiensmódszerrel, a (3-8) egyenlettel határozhatjuk meg. A μ lépésmagyságot optimális (maximális) értékét a

$$\frac{\partial J(\mathbf{h}, \mu)}{\partial \mu} = 0 \quad (3-14)$$

egyenlethez meg lehet határozni, mivel ismert a gradiens,

$$\mu_{\text{opt}} = \mu_{\text{max}} = \frac{2}{\mathbf{x}^T \mathbf{x}} = \frac{2}{\text{bemeneti jelteljesítmény}} \quad (3-15)$$

adódik. Ha a jelekről a priori semmit nem tudunk, akkor a paraboloid és a gradiens is ismeretlen, az LMS algoritmus

$$0 < \mu < \mu_{\text{max}} \quad (3-16)$$

lépésmagyság esetén stabilis.

3. Konvergencia. A (3-11) szerinti közelítés meglehetősen durva, ezért egyrészt kvalitatíve nyilvánvaló, hogy a Wiener-Hopf egyenletnek csak közelítő megoldását kapjuk, másrészt annyira távol kerültünk ezáltal a Wiener-problémától, hogy feltehető a kérdés, mi történik a Wiener-feladat által le nem fedett esetekben, nemstacionárius jeleknél.

Az első kérdésre számszerű választ keresve, írjuk fel az átlagos négyzetes hiba Wiener-minimumát, (3-2)-be helyettesítve (3-5)-öt:

$$J(\mathbf{h}_{\text{opt}}) = J_{\text{min}} = \sigma_{d|1}^2 - \mathbf{p}^T \mathbf{R}^{-1} \mathbf{p} = \sigma_{d|1}^2 - \mathbf{p}^T \mathbf{h}_{\text{opt}} \quad (3-17)$$

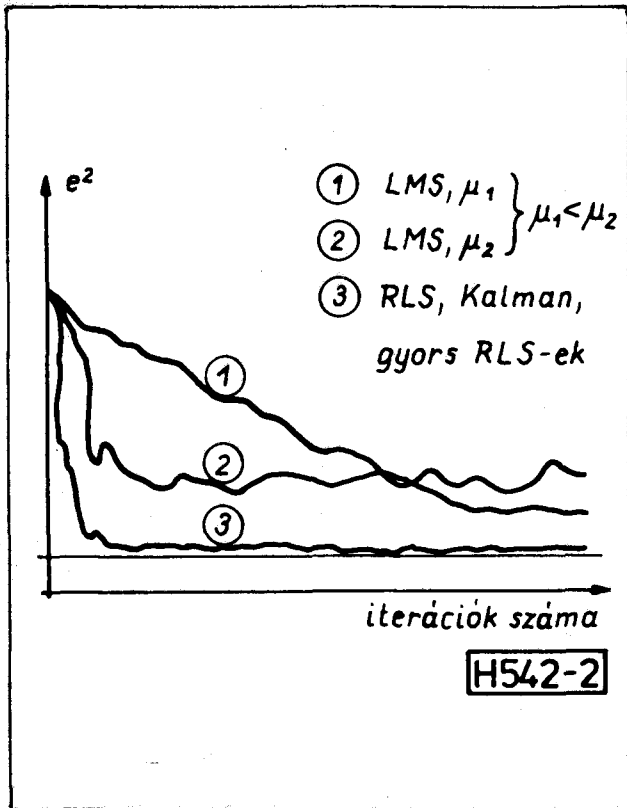
Belátható, hogy a maradékhiba végértékének várható értéke nem nulla:

$$E\{J_{\text{ex}}(\infty)\} = E\left\{\lim_{k \rightarrow \infty} (J(\mathbf{h}_k) - J_{\text{min}})\right\} = \frac{\mu J_{\text{min}} \text{tr } \mathbf{R}}{2 - \mu \text{tr } \mathbf{R}} \quad (3-18)$$

Ez azt jelenti, hogy az algoritmus a μ lépésmagyságtól függő sebességgel és maradékhibával, mindig a Wiener-optimumhoz tartozónál nagyobb maradékhibához konvergál négyzetes átlagban, ha a (3-16) feltételt betartjuk. Az is belátható továbbá, hogy (3-16) esetén

$$E\left\{\lim_{k \rightarrow \infty} \mathbf{h}_k\right\} = \mathbf{h}_{\text{opt}}, \quad (3-19)$$

vagyis a szűrőegytárhóvektor sokaság-átlagban a Wiener-optimumhoz konvergál. Ha μ értéke nagy, akkor a konvergencia gyorsabb, de nagyobb a maradékhiba, míg ha kicsi, akkor a konvergencia lassabb, de kisebb a maradékhiba. Konvergencia esetén a szűrő egytárhóvektora a Wiener-optimum körül μ által meghatározott amplitúdóval véletlenszerűen oszcillál. A konvergenciátranziensek tipikus menete a 2. ábrán látható.



2. ábra Adaptív szűrők tipikus konvergenciátranziensei

4. A 3. pontban említett második kérdésre választ keresve, nemstacionárius esetben (nemstacionárius jelek vagy idővariáns rendszer modellezése) μ értékét alulról az a követelmény korlátozza, hogy az időben változó paraboloidot követni kell, felülről pedig az előírt maradékhiba. Így rendszerint μ -t e két hatás kompromisszumaként kell megválasztani, ami általában tapasztalaton alapszik.

5. Kondicionáltság. A konvergencia sebessége nem csupán a μ értékétől függ, hanem a bemeneti jeltől is. Mint az a gradiensmódszer esetében jól ismert, ha az R korrelációs mátrix rosszul kondicionált, vagyis legnagyobb és legkisebb sajátértékének aránya nagy, akkor a hibafelület-paraboloid metszete erősen ellipszissé válik és így a konvergencia lassú lesz, sebességét döntően a legkisebb sajátérték határozza meg. Ez lényeges hátránya az LMS algoritmusnak.

6. A stochasztikus approximációval összevetve, az LMS algoritmus esetén a μ lépésmagyság állandó a konvergencia során, ami véges memóriának (felejtő algoritmusnak) felel meg. Emiatt az LMS algoritmus nem csak statikus környezetben alkalmazható, amint a 4. pontban vázoltuk.

7. Véges szóhosszúság hatása [1], [3]. A digitális megvalósítás következtében mind a bemeneti jelek, mind a belső számábrázolás véges pontosságúak. Az algoritmus rekurzív jellege a hiba folyamatos akkumulálódásához, az pedig numerikus instabilitáshoz vezethet. Numerikusan stabil algoritmus esetén a hibák akkumulálódása véges

határokon belül megy végbe. Részletes analízisbe nem bocsátkozhatunk, csak azt említjük meg, hogy az LMS algoritmus általában numerikusan stabil, pontossága állandósult állapotban viszont függ a μ lépésmagyságtól és a bemeneti jel kondicionáltságától. Nevezetesen, a kvantálási hiba μ csökkenésével egy bizonyos határ elérése után nő. Stacionárius esetre is található tehát egy optimális μ , mely az összhibát (maradékhiba + kvantálási hiba) minimalizálja, de ez többnyire ismeretlen.

8. Megemlítjük, hogy az LMS algoritmusnak számos módosított változata van [5], [6], [9], [10], [7]. A módosítások egyrészt a konvergencia stabilizálását (pl. jelteljesítményre normalizált változat), másrészt az algoritmus gyorsítását célozzák. Ez utóbbira példa az "előjeles" algoritmus, ahol a (3-13) rekurzióban x_k helyett csak annak előjele szerepel, felére csökkentve ezáltal a műveletigényt, valamint a transzformált tartományban végzett adaptáció. A frekvenciatartományban adaptálva, szokásos szűrőhosszak esetén egy-két nagyságrenddel kisebb lehet az adaptáció műveletigénye. A Karhunen-Loève transzformáció azáltal javítja a konvergenciát, hogy az adaptáció sebessége függetlenné válik a korrelációs mátrix kondicionáltságától [7].

4. Kalman-szűrő

Az előző pontban ismertetett LMS algoritmus fő hátránya az, hogy nem használja fel az adaptációhoz az összes rendelkezésre álló információt a bemeneti jelről s emiatt lassú a konvergenciája. A továbblépést a rekurzív becslőn alapuló algoritmusok kidolgozása jelenti. Induljunk ki abból, hogy az y_1, y_2, \dots, y_{n-1} véletlen változókból már ismerjük az x_{n-1} mennyiség legkisebb átlagos négyzetes hibájú becslőjét, $x_{n-1}(Y_{n-1})$ -et, utalva e jelölésben arra, hogy a becslést az első $n-1$ számú y_k adta ($y_k=0, k < 0$). A feladat x_n becslőjének, $x_n(n)$ meghatározása $x_{n-1}(Y_{n-1})$ és y_n ismeretében. Már itt megemlítjük, hogy a FIR szűrőkre levezetett Kalman-algoritmus alakilag az RLS (recursive least squares) algoritmussal azonos, kapcsolatukra az 5. pontban világitunk rá.

A Kalman-szűrés általánosabb feladata az alábbi [5], [6]. Egy lineáris, dinamikus, diszkrét idejű rendszer leírására az

$$x_{n+1} = \Phi_{n+1,n} x_n + v_n \quad (4-1)$$

$$y_n = C_n x_n + u_n \quad (4-2)$$

állapotegyenletet és mérési egyenletet alkalmazunk, ahol x_n az állapotvektor, y_n a rendszer kimenete; a $\Phi_{n+1,n}$ állapotátmeneti mátrix - mely a rendszer n és $n+1$ időpontbeli állapotait kapcsolja össze - és a C_n mátrix ismertek, valamint ismert a v_n és u_n független, zérus középértékű fehér zaj Q_v ü. Q_u korrelációs

mátrixa. A feladat az y_1, y_2, \dots, y_n megfigyelések birtokában meghatározni az x_k állapotvektor legkisebb átlagos négyzetes hibájú becslőjét.

Jelölje $\hat{y}_n(Y_{n-1})$ az y_n változónak az összes addigi, y_1, y_2, \dots, y_{n-1} adatokból predikált értékét. Az $\hat{y}_n(Y_{n-1})$ mennyiséghez képest az y_n valódi érték pontosan annyi új információt jelent, mint az

$$\alpha_n = y_n - \hat{y}_n(Y_{n-1}), n=1, 2, \dots \quad (4-3)$$

predikációs hiba vagy innovation (Kailath). Miután v_n és u_n korrelálatlan, beszorozva a (4-1) egyenletet mindkét oldalát u_n^T -tal és várható értéket véve

$$E\{x_k u_n^T\} = 0, k, n \geq 0 \quad (4-4)$$

adódik, (4-2) felhasználásával pedig

$$E\{y_k u_n^T\} = 0, 0 \leq k \leq n-1$$

$$E\{y_k v_n^T\} = 0, 0 \leq k \leq n. \quad (4-5)$$

A (4-2) egyenlet alapján

$$\hat{y}_n(Y_{n-1}) = C_n \hat{x}_n(Y_{n-1}) + \hat{u}_n(Y_{n-1}) \quad (4-6)$$

írható, de (4-5) miatt (4-6) utolsó tagja zérus, vagyis a predikációs hiba (4-3) szerint kifejezhető

$$\alpha_n = y_n - \hat{x}_n C_n(Y_{n-1}) \quad (4-7)$$

alakban. Jelölje α_n korrelációs mátrixát

$$\Sigma_n = E\{\alpha_n \alpha_n^T\} \quad (4-8)$$

Az ortogonalitás elve miatt

$$E\{[\hat{x}_i - \hat{x}_i(Y_n)] \alpha_k^T\} = 0, k=1, 2, \dots, n \quad (4-9)$$

vagyis annak érdekében, hogy a becslőt α_k -k lineáris kombinációjaként

$$\hat{x}_1(Y_n) = \sum_{k=1}^n B_{1k} \alpha_k \quad (4-10)$$

alakban írassuk fel,

$$E\{x_i \alpha_k^T\} = B_{ik} \Sigma_k \quad (4-11)$$

következtetjük az előző két egyenletből, így tehát

$$\hat{x}_i(Y_n) = \sum_{k=1}^n E\{x_i \alpha_k^T\} \Sigma_k^{-1} \alpha_k \quad (4-12)$$

adódik. A (4-1) egyenletet beszorozva jobbról α_k^T -val és várható értéket képezve, (4-5) miatt

$$\begin{aligned} E\{x_{n+1} \alpha_k^T\} &= E\{[\phi_{n+1,n} x_n + v_n] \alpha_k^T\} = \\ &= \phi_{n+1,n} E\{x_n \alpha_k^T\}. \end{aligned} \quad (4-13)$$

Ebből következően

$$\sum_{k=1}^{n-1} E\{x_{n+1} \alpha_k^T\} \Sigma_k^{-1} \alpha_k = \phi_{n+1,n} \sum_{k=1}^{n-1} E\{x_n \alpha_k^T\} \Sigma_k^{-1} \alpha_k \quad (4-14)$$

Alkalmazva (4-12)-t $i=n+1$ -re (4-14) bal oldalán és $i=n$ -re a jobb oldalán (n helyére $n-1$ -et írva),

$$\hat{x}_{n+1}(Y_n) = \phi_{n+1,n} \hat{x}_n(Y_{n-1}) + G_n \alpha_n \quad (4-15)$$

adódik, ahol

$$G_n = E\{x_{n+1} \alpha_n^T\} \Sigma_n^{-1} \quad (4-16)$$

az ún. Kalman-erősítés.

A (4-15) egyenlet az állapotvektor $n+1$ -edik időpontban vett, első n adatból predikált értékére ad egy rekuziót. A szűrési probléma viszont $\hat{x}_n(Y_n)$ meghatározását jelenti. Ehhez a (4-1) egyenletben írjuk fel a változókat, mint legkisebb átlagos négyzetes hibájú becslőket; (4-5) figyelembevételével így

$$\hat{x}_{n+1}(Y_n) = \phi_{n+1,n} \hat{x}_n(Y_n) \quad (4-17)$$

adódik. Az állapotátmeneti mátrix definíciója szerint tehát

$$\hat{x}_{n+1}(Y_n) = \phi_{n+1,n}^{-1} \hat{x}_{n+1}(Y_n) = \phi_{n,n+1} \hat{x}_{n+1}(Y_n) \quad (4-18)$$

(4-15) felhasználásával

$$\begin{aligned} \hat{x}_{n+1}(Y_n) &= \hat{x}_n(Y_{n-1}) + \phi_{n,n+1} G_n \alpha_n = \\ &= \phi_{n,n-1} \hat{x}_{n-1}(Y_{n-1}) + \phi_{n,n+1} G_n \alpha_n \end{aligned} \quad (4-19)$$

kapjuk és (4-7) is kifejezhető

$$\alpha_n = y_n - C_n \phi_{n,n-1} \hat{x}_{n-1}(Y_{n-1}) \quad (4-20)$$

alakban. A (4-19) és (4-20) egyenlet írja le az optimális lineáris szűrőt, a Kalman-szűrőt, melynek jelfolyamábrája a 3. ábrán látható. A bemenet y_n , a kimenet $\hat{x}_n(Y_n)$. A (4-16) egyenlettel definiált Kalman-erősítés számítására is megadható egy rekuzív formula, mely az algoritmust teljessé teszi. Ennek levezetésétől eltekintünk, a végeredmény:

$$\begin{aligned} G_n &= \phi_{n+1,n} K_{n,n-1} C_n^T [C_n K_{n,n-1} C_n^T + Q_{un}]^{-1} \\ K_{n+1,n} &= \phi_{n+1,n} K_n \phi_{n+1,n}^T + Q_{vn} \end{aligned} \quad (4-21)$$

$$K_n = K_{n,n-1} - \phi_{n,n+1} G_n C_n K_{n,n-1}$$

Kezdeti feltételek:

$$\hat{x}_1(0) = 0$$

$$K_1(0) = E\{x_1 x_1^T\}. \quad (4-22)$$

A Kalman-szűrőt adaptív FIR szűrő algoritmus kialakítására többen is alkalmazták, az alábbiakban Godard módszerét követjük.

Először a stacionárius jeleket feldolgozó adaptív FIR szűrővel foglalkozunk (1. ábra). Ennek optimális együttthatóvektorát a Wiener-Hopf egyenlet definiálja, a stacionaritás következtében

$$h_{n+1} = h_n (= h_{opt}). \quad (4-23)$$

Ez felel meg a (4-1) állapotegyenletnek,

$$d_n = x_n^T h_n + e_n \quad (4-24)$$

pedig a mérési egyenletnek. A Kalman-szűrő elméletét alkalmazva erre a speciális estre, az alábbiak adódnak:

$$(4-19) \rightarrow \hat{h}_n = \hat{h}_{n-1} + g_n \alpha_n \quad (4-25)$$

$$(4-20) \rightarrow \alpha_n = d_n - x_n^T \hat{h}_{n-1} \quad (4-26)$$

$$(4-21) \rightarrow g_n = K_{n-1} x_n [x_n^T K_{n-1} x_n + J_{min}]^{-1} \quad (4-27)$$

$$K_n = K_{n-1} - g_n x_n^T K_{n-1} \quad (4-28)$$

A (4-27) egyenletben az e_n hiba optimális értékéhez tartozó varianciáját J_{min} -nel jelöltük, ez a Wiener-megoldás hibaértéke (l. a (3-17) egyenlet).

Kezdeti feltételek (4-22) szerint

$$\hat{h}_0 = 0, K_0 = E\{h_0 h_0^T\} = cI, c > 0 \quad (4-29)$$

Az algoritmus megkívánja J_{min} a priori ismeretét. Mivel ez a gyakorlatban ismeretlen, az alábbi becsléssel szokás élni:

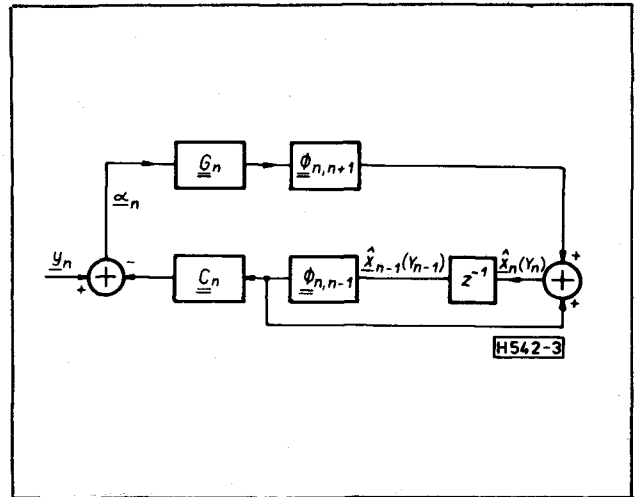
$$J_{min} \sim \frac{\sigma_d^2}{100} \dots \frac{\sigma_d^2}{1000} \quad (4-30)$$

A (4-25)-(4-30) egyenletekkel definiált Kalman-algoritmus főbb tulajdonságai az alábbiak:

1. Az algoritmus a Wiener-optimumhoz konvergál ergodikusság esetén:

$$\lim_{n \rightarrow \infty} \hat{h}_n = h_{opt} \quad (4-31)$$

Emiatt a Kalman-algoritmus maradékhibája nulla. Az LMS algoritmusnál viszont nem maga az együttthatóvektor, hanem annak csak a sokaság szerinti átlagértéke konvergál az optimumhoz.



3. ábra A Kalman-szűrő jelfolyamábrája

2. A Kalman-szűrő a becsléshez az összes addigi információt felhasználja az LMS algoritmussal ellentétben, aminek az a következménye, hogy konvergenciája tipikusan egy nagyságrenddel gyorsabb (l. a 2. ábrát).

3. A Kalman-algoritmus elvileg mindig stabil, az LMS algoritmus nem.

4. A kedvező tulajdonságok ára a lényegesen nagyobb, $O(N^2)$ számításgény a mátrix-vektor szorzások miatt.

A Kalman-szűrő második alkalmazása adaptív FIR szűrőkre a nemstacionárius jelek esete. A nemstacionaritás arra vezet, hogy az együttthatóvektor időben véletlenszerűen változik:

$$h_{n+1} = h_n + v_n \quad (4-32)$$

$$E\{v_n\} = 0, E\{v_n v_n^T\} = Q_v \quad (4-33)$$

A mérési egyenlet

$$d_n = x_n^T h_n + e_n \quad (4-34)$$

A Kalman-algoritmus ekkor az alábbi:

$$(4-19) \rightarrow \hat{h}_n = \hat{h}_{n-1} + g_n \alpha_n \quad (4-35)$$

$$(4-20) \rightarrow \alpha_n = d_n - x_n^T \hat{h}_{n-1} \quad (4-36)$$

$$(4-21) \rightarrow g_n = K_{n-1} x_n [x_n^T K_{n-1} x_n + \frac{J_{min}}{1+q}]^{-1} \quad (4-37)$$

$$K_n = (1+q)[K_{n-1} - g_n x_n^T K_{n-1}] \quad (4-38)$$

ahol Q_v értékére szükség volt, ezért feltételeztük, hogy $Q_v = qK_n$. Kis q esetén lassú a konvergencia, jó a követés, nagy q esetén fordítva van. A Q_v mátrix ilyen

közelítése tapasztalat szerint általában kielégítő eredményekre vezet viszonylag lassan változó környezet esetén.

A Kalman-szűrő minimális varianciájú becslő, mert a (4-28) ill. (4-38) egyenletben álló

$$K_n = E\{(x_n - \hat{x}_n(Y_n))(x_n - \hat{x}_n(Y_n))^T\}$$

mátrix nyomát minimalizálja. Gauss-folyamatok esetén az optimális szűrőt, míg nem Gauss-folyamatok esetén az optimális lineáris becslőt adja.

IRODALOM

[1] Caraiscos, C. - Liu, B.: A roundoff error analysis of the LMS adaptive algorithm. IEEE Trans. ASSP, vol. ASSP-32, No. 1, Feb. 1984, pp. 34-41.

[2] Carayannis, G. - Manolakis, D.G. - Kalouptsidis, N.: A fast sequential algorithm for adaptive filtering. IEEE Trans. ASSP, vol. ASSP-31, No. 6, Dec. 1983, pp. 1394-1402.

[3] Cioffi, J.M.: Limited-precision effects in adaptive filtering. IEEE Trans. CAS, vol. CAS-34, No. 7, July 1987, pp. 821-833.

[4] Cioffi, J.M. - Kailath, T.: Fast, recursive-least-squares transversal filters for adaptive filtering. IEEE Trans. ASSP, vol. ASSP-32, No.2, Apr. 1984, pp. 304-337.

[5] Cowan, C.F.N. - Grant, P.M.: Adaptive filters. Prentice-Hall, Engl. Cliffs, N.J., 1985.

[6] Haykin, S.: Adaptive filter theory. Prentice-Hall, Engl. Cliffs, N.J., 1986.

[7] Narayan, S.S. - Peterson, A.M. - Narashimha, M.J.: Transform domain LMS algorithm. IEEE Trans. ASSP, vol. ASSP-31, No. 3, June 1983, pp. 609-614.

[8] Robbins, H. - Monro, S.: A stochastic approximation method. Ann. Math. Statis., vol. 22, 1951, pp. 400-407.

[9] Treichler, J.R. - Johnson, C.R. - Larimore, M.G.: Theory and design of adaptive filters. J.Wiley, N.J., 1987.

[10] Widrow, B. - Stearns, S.: Adaptive signal processing. Prentice-Hall, Engl. Cliffs, N.Y., 1985.