

Projektkurs i adaptiv signalbehandling

SKALÄRA WIENERFILTER

Wienerfiltrering syftar till att optimera linjära filter så att estimeringsfel minimeras i medelkvadratmening. Filterkonstruktionen utgår från att linjära stokastiska modeller av signaler och av brus är kända. Wienerfiltrering och Kalmanfiltrering är nära besläktade: Båda minimerar estimeringsfel i kvadrat för signaler beskrivna av linjära system. Kalmanfilter konstrueras utifrån tillståndsmodeller som kan ha tidsvariabel dynamik. Wienerfiltrering utgår för det mesta från *tidsinvarianta modeller* på *insignal-utsignalform*, som överföringsfunktioner och ARMA-modeller. Wienerlösningar är intressanta att studera bland annat därför att de indikerar ideala konvergenspunkter hos adaptiva filter.

Klassisk Wienerfiltrering [5, 16] baseras på att man manipulerar frekvensfunktioner. Se t.ex Kap. 7.3 i Hayes bok [8]. Här ges en alternativ förklaring av Wienerfiltrering som gör det lättare att förstå funktionen hos kausala IIR-Wienerfilter. Framställningen baseras väsentligen på artiklarna [1, 2] och bok-kapitlet [3].

1 Filtrering, smoothing, prediktion, avfaltung

För att demonstrera hur Wienerfilter härleds och fungerar så utgår vi från en förenklad variant av de problem som studeras i [3], se Figur 1. En tidsdiskret skalär signal $u(k)$, ska skattas. Signalen modelleras som en ARMA-process med känt spektrum. Den kan ha gått genom en känt dynamiskt system, som vi här låter vara en stabil IIR-kanal. Kanalsignalen störs sedan av ett ARMA-brus $w(k)$.

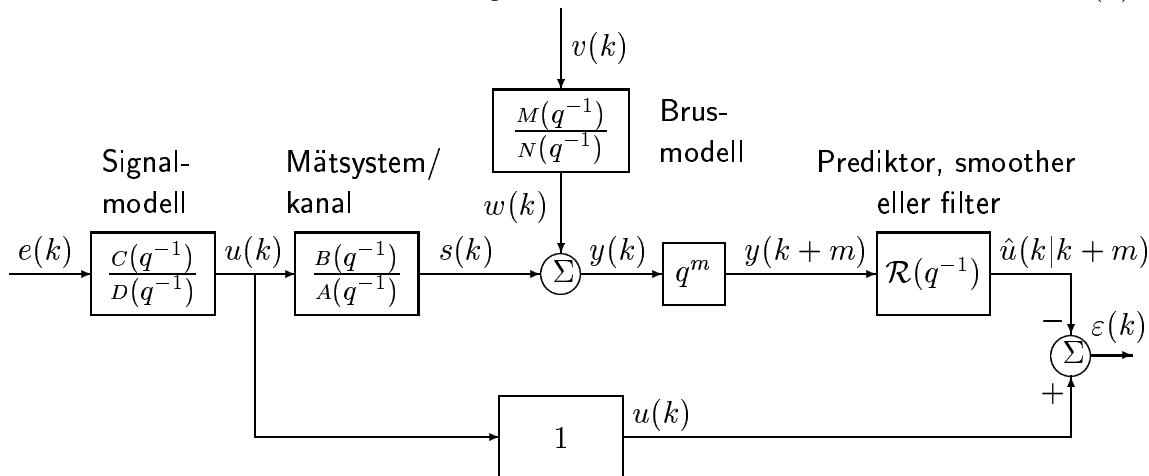


Figure 1: Ett skalärt Wienerfiltreringsproblem. Sekvensen $\{u(k)\}$ ska estimeras baserad på mätdata $\{y(k)\}$, insamlade till och med tidpunkten $k + m$. Om $B(q^{-1})/A(q^{-1}) \neq q^{-m}$, så utgör problemet ett avfaltungsproblem (deconvolution).

Signal och brus antas inbördes okorrelerade, dvs $Ee(k)v^*(j) = 0 \forall k, j$.¹ Model-
lens "drivande brus" $e(k)$ och $v(k)$ antas också vara vita, ha kända varianser λ_e
respektive λ_v samt ha medelvärden noll². Alla signaler får vara komplexvärda.

Ett linjärt tidsinvariant filter

$$\hat{u}(k|k+m) = \mathcal{R}(q^{-1})y(k+m) \quad (1)$$

ska härledas som utnyttjar mätdata $y(k+m)$ och som minimerar medelkvadratfelet

$$J = E|\varepsilon(k)|^2 \quad (2)$$

$$\varepsilon(k) = u(k) - \hat{u}(k|k+m) \quad (3)$$

Här är $\mathcal{R}(q^{-1})$ en rationell överföringsoperator, d.v.s. en rationell funktion i bakåt-
skiftoperatorn q^{-1} , där $q^{-1}y(k) = y(k-1)$. Det är möjligt att ta fram en enda
uppsättning designekvationer för tre typer av signalestimatorer:

- $m < 0$: **Prediktorer**, som förutsäger $u(k)$. Prediktion fungerar sämre ju längre
prediktionshorisont $-m$ man väljer och ju mera brus som påverkar mätningen.
- $m = 0$: **Filter**, som utnyttjar det senaste datavärdet. I detta fall uppnås bättre
prestanda än vid prediktion. Prestanda avgörs av egenskaperna hos kanalen
och av brusnivån. För problem med dynamiska kanaler krävs en *avfaltung*
(*deconvolution*). Idealt skulle då kanalen behöva inverteras exakt, vilket dock
är omöjligt om kanalen saknar en stabil invers (är icke-minfas).
- $m > 0$: **Fixed-lag smoothers**, som utnyttjar "framtida" data för skattning av $u(k)$.
Om möjligt ska man utnyttja smoothing! Framtida mätvärden innehåller i
allmänhet viktig information som kan minska estimeringsfelet jämfört med
filtrering. Resultaten blir allt bättre ju längre smoothing-lag m som väljs.
Man når dock en gräns där ytterligare ökning av m bara ger små förbättringar³.
Om $m \rightarrow \infty$ så erhålls det *ickekausala wienerfiltret*.

Ibland kan man behöva utnyttja generellare problemställningar, där t.ex. blocket "1"
i Figur 1 ersätts med ett filter. Felet $\varepsilon(k)$ kan även viktas olika vid olika frekvenser
och signalerna kan få utgöra kolonnvektorer istället för skalärer.

För att säkerställa att estimatorn bara använder data från de tidpunkter vi bestämt,
så måste själva filtet $\mathcal{R}(q^{-1})$ vara kausalt. (Detta gäller även då smoothing ut-
nyttjas). Dessutom måste vi kräva att $\mathcal{R}(z^{-1})$ har alla poler strikt innanför
 $|z| = 1$ d.v.s. att Wienerestimatorn är stabil. Om estimatorn innehöll gömda

¹Man kan alltid göra en uppdelning där signal och brus är inbördes okorrelerade. Signal-
korrelerade komponenter hos bruset kan alltid separeras ut ur brusbeskrivningen.

²Om signal eller brus har medelvärden skilda från noll kan dessa medelvärden hanteras sep-
arat. Det är dock bekvämare att i sådana fall införa integrerande ARMA-modeller, som ger
Wienerfiltret lämpliga egenskaper. Man kan visa att om $D(z^{-1}) = 0$ för $z = 1$ så kommer filtret
att få rätt statisk förstärkning. Se sid 406-407 i [3]. Om $N(z^{-1}) = 0$ för $z = 1$ så kommer filtret
att få ett nollställe vid $z = 1$, så statistiska störningsnivåer släpps ej igenom filtret.

³Det finns även en variant "fixed interval smoothing" där man vid offline-filtrering använder
en hel datasekvens för att skatta värden inne i sekvensen. Sådana estimatorer implementeras
bäst på Kalmanform, se [4].

instabila moder så skulle inverkan av avrundningsfel växa katastrofalt. Under förutsättningarna ovan kommer estimeringsfelet att vara en stationär process med medelvärde noll.

Minimering av variansen hos estimeringsfelet

$$\varepsilon(k) = \left(1 - \mathcal{R}(q^{-1})q^m \frac{B(q^{-1})}{A(q^{-1})}\right) u(k) - \mathcal{R}(q^{-1})w(k+m) \quad (4)$$

utgör en kompromiss mellan förmågan hos estimatorn att skatta $u(k)$ (den första termen) och dess brusförstärkning i den sista termen.

Man kan konstruera Wienerestimatorer på två sätt:

1. Ett linjärt filter med förutbestämd struktur och ordning ansätts och dess koefficienter justeras så att kriteriet (2) minimeras. Detta kan ske med en numerisk sökmethode. Om $\mathcal{R}$ ansätts som ett *FIR-filter* har vi ett *entydigt minimum* och lösningen kan erhållas på sluten form via ett linjärt ekvations-system. Se Kap. 7.2 i Hayes bok [8]. Om man ansätter IIR-filter av alltför låg ordning så kan tyvärr flera lokala minima förekomma.
2. Optimala IIR-filter av korrekt ordning kan dock bestämmas entydigt genom en analytisk beräkning av det *kausala IIR-Wienerfiltet*.

Vi ska här studera beräkning av och egenskaper hos kausala IIR-Wienerfilter.

Notation. Linjära tidsinvarianta modeller och filter betecknas med script-bokstäver såsom $\mathcal{F}$ eller $\mathcal{R}$. Om de parametreras av polynom i bakåt-skiftoperatorn q^{-1} eller variabeln z^{-1} så förenklas härledning och lösningen av designekvationerna.⁴ I tidsdomänen används i bakåtskift-operatorn och polynom av grad np ges av

$$P(q^{-1}) = p_0 + p_1q^{-1} + \dots + p_{np}q^{-np} \quad .$$

Om $p_0 = 1$ sägs polynomet vara *moniskt*. Med framåtskift-operatorn q kan man då definiera det *reciproka polynomet*

$$P_*(q) \triangleq p_0^* + p_1^*q + \dots + p_{np}^*q^{np} \quad ,$$

där $*$ representerar komplexkonjugering, då vi tillåter modeller med komplexa koefficienter. Argumenten q^{-1} eller q kommer ofta att utelämnas för att snygga upp uttrycken, där de utan risk kan underförstås. Vi kallar polynom för *stabila* om de kan utgöra nämnare för stabila system, dvs om alla nollställena hos $z^{np}P(z^{-1})$ ligger innanför enhetscirkeln $|z| = 1$. Observera att nollställena hos $P_*(z)$ kommer att vara speglade i enhetscirkeln, jämfört med nollställena hos $z^{np}P(z^{-1})$. För stabila $P(z^{-1})$ har $P_*(z)$ alla nollställena *utanför* $|z| = 1$.

⁴Till skillnad från amerikansk litteratur skiljer vi här på *tidsdomän-operatorn* q och den *komplexa variabeln* z , även om ekvationer sällan ändras då man byter den ena notationen mot den andra.

2 En variationsmetod för härledning

Kausala IIR-Wienerfilter kan härledas på flera sätt:

- Med en ”klassisk” härledning som i Kap. 7.3 i [8].
- Genom att derivera kriteriet (2) med avseende på filtrets koefficienter och sätta derivatorna till noll.
- Genom att utveckla kriteriet (2) och kvadratkomplettera tills filtret kan väljas så att en term hos J elimineras helt, medan övriga termer inte kan påverkas av filtret [6].
- Genom en variationsmetod, där man konstruerar filtret så att det inte låter sig förbättras.

Alla vägar leder fram till samma mål (se avsnitt III i [3]) men variationsmetoden är den enklaste, och beskrivs här. Metodens princip är att man inför en potentiell additiv förändring (variation) $\nu(k)$ av estimatet (1)

$$\tilde{u}(k|k+m) \triangleq \hat{u}(k|k+m) + \nu(k) = \mathcal{R}(q^{-1})y(t+m) + \mathcal{T}(q^{-1})y(t+m) \quad , \quad (5)$$

se Figur 2. Eftersom en grundförutsättning är att vi ska utföra stabil linjär tidsinvariant filtrering baserad på mätsignaler upp till och med tiden $k+m$, så måste variationen kunna uttryckas som ovan, där $\mathcal{T}(q^{-1})$ är någon stabil, kausal och rationell överföringsoperator. Om $y(k)$ är en stationär signal så blir $\nu(k)$ därför också stationär.⁵ Variansen hos det modifierade estimatets fel $\tilde{\varepsilon}(k) = u(k) - \tilde{u}(k|k+m)$ blir nu

$$\begin{aligned} E|\tilde{\varepsilon}|^2 &= E(u - \hat{u} - \nu)(u - \hat{u} - \nu)^* \\ &= E(u - \hat{u})(u - \hat{u})^* - E(u - \hat{u})\nu^* - E\nu(u - \hat{u})^* + E\nu\nu^* \quad . \end{aligned}$$

Kan estimatorn konstrueras så att ingen variationssignal $\nu(k)$ kan förbättra estimatet i medelkvadratmening? Det ursprungliga estimeringsfelet är $\varepsilon(k) = u(k) - \hat{u}(k)$. Man ser att $E|\tilde{\varepsilon}(k)|^2 \geq E|\varepsilon(k)|^2$, med likhet för $\nu(k) = 0$ om och endast de potentiellt negativa bidragen från korstermerna helt elimineras, dvs då

$$E\varepsilon(k)\nu^*(k) = E\nu(k)\varepsilon^*(k) = 0 \quad .$$

Med andra ord är estimatorn optimal om och endast om estimeringsfelet är okorrelerat med alla tillåtna variationssignaler. Eftersom korrelation i signalernas Hilbert-rum definierar skalärprodukten så betyder detta att

Estimeringsfelet skall vara ortogonalt mot alla tillåtna variationer.

⁵Om $y(k)$ inte är stationär, därför att vi har system med poler på enhetscirkeln, så måste $\nu(k)$ ändå särkerställas som stationär, med medelvärde noll. Då måste vi kräva och anta att täljaren i $\mathcal{T}(q^{-1})$ innehåller nollställena på eller utanför enhetscirkeln som eliminerar polerna i systemet som genererar $y(t+m)$ som gör att denna signal får obegränsad varians. (Detta förutsätter naturligtvis att dessa polers lägen är exakt kända, vilket i stort sett bara är realistiskt för integratorer, dvs poler i $z = 1$.)

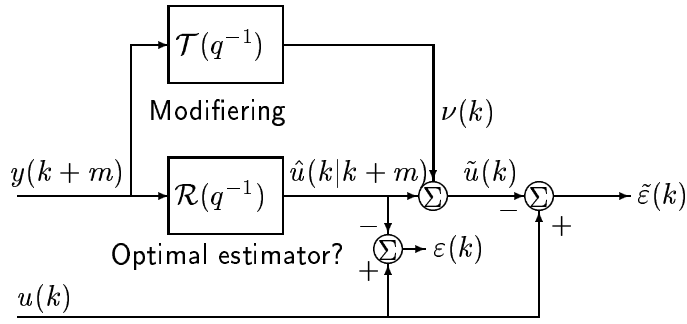


Figure 2: Om ett signal-estimat $\hat{u}(k|k+m)$ inte kan förbättras av någon tillåten additiv variation $\nu(k)$, så är estimatorn $\mathcal{R}(q^{-1})$ optimal.

Ortogonalitetsvillkoret innebär att

$$E\varepsilon(k)y(k+n) = 0 \quad n = m, m-1, \dots,$$

eftersom filtret $\mathcal{T}(q^{-1})$ kan ha godtyckliga koefficienter. Vi får alltså oändligt många villkor i tidsdomänen vid optimering av IIR-filer. För att undvika denna komplikation föredrar man att hantera ortogonalitetsvillkor i frekvensdomänen.

Det finns en systematisk teknik för att härleda Wienerfilter för en given problemformulering, genom att säkerställa ortogonaliteten.

1. Uttryck mätsignalen som en ekvivalent ARMA-modell, en *innovationsmodell* driven av vitt brus (innovationssekvensen). Wienerfiltret kommer för de flesta problem att representeras som två seriekopplade filter, där inversen av innovationsmodellen utgör det första filtret. Beräkning av täljarpolynomet hos en innovationsmodell sker genom en *spektralfaktorisering*.
2. Inför en variationsterm och uttryck villkoret $E\varepsilon(k)\nu^*(k) = 0$ i frekvensdomänen, via korspektrum $\phi_{\varepsilon\nu^*}$ ⁶ och Parsevals formel

$$E\varepsilon(k)\nu(k)^* = \frac{1}{2\pi j} \oint_{|z|=1} \boldsymbol{\phi}_{\varepsilon\nu^*} \frac{dz}{z} . \quad (6)$$

3. Säkerställ ortogonaliteten genom att eliminera *alla* poler hos integranden $\boldsymbol{\phi}_{\varepsilon\nu^*}$ innanför integrationsvägen, som är enhetscirkeln. Då blir kurvintegralen noll. Detta villkor kommer att motsvara en eller flera *Diophantiska ekvationer*. Det optimala Wienerfiltret är nu entydigt specificerat.⁷
4. Om modellerna ej är stabila måste man som ett sista steg verifiera att estimeringsfelet $\varepsilon(k)$ verkligen blir en stationär signal med ändlig varians och medelvärde noll, även då detta *inte* gäller för alla andra signaler. Se Appendix B i [3]. För stabila modeller är detta steg onödigt.

Metodiken kan enkelt generaliseras till flervariabla system, där (6) blir en matris.

⁶Om $g(k) = q^{-g}\mathcal{G}(q^{-1})e(k)$ och $f(k) = q^{-f}\mathcal{F}(q^{-1})e(k)$, med $e(k)$ vit, så är med våra beteckningar $\phi_{gf^*} = z^{-g}\mathcal{G}(z^{-1})\lambda_e z^f \mathcal{F}_*(z)$.

⁷Man kan visa att några andra möjligheter ej kan förekomma, t.ex att integranden har poler innanför enhetscirkeln men att alla residuer tar ut varandra.

3 Wienerestimatorns struktur och funktion

Vi tillämpar nu metodiken som skisserades i förra avsnittet för att härleda en lösning på Wienerproblemet i Figur 1. Låt oss införa

Antagande A. ARMA-modellerna för den brusfria signalen $s(k) = (BC/AD)e(k)$ och bruset $w(k) = (M/N)v(k)$ är stabila⁸ och kausala och antas ej ha gemensamma nollställen på enhetscirkeln.

De två designekvationerna för estimatorn $\mathcal{R}$ som minimerar (2) härleds i Appendix A. Man behöver lösa spektralfaktoriserings-ekvationen

$$r\beta\beta_* = CC_*BB_*NN_* + \rho MM_*AA_*DD_* \quad (7)$$

som entydigt ger ett stabilt moniskt polynom $\beta(q^{-1})$ och en skalär r , om modellen och brusvarianskvoten

$$\rho = \lambda_v/\lambda_e$$

är given. Vi beräknar även två polynom $Q_1(q^{-1})$ och $L_*(q)$ ur en Diofantisk ekvation,⁹ som är en linjär polynomekvation i de två obekanta

$$q^{-m}CC_*B_*N_* = r\beta_*Q_1 + qDL_* \quad . \quad (8)$$

Wienerestimatorn ges då av

$$\hat{u}(k|k+m) = \frac{Q_1NA}{\beta}y(k+m) \quad . \quad (9)$$

Man kan notera att brusmodellens nämnarpolynom N ingår i estimatorns täljare. Resonanstoppar hos bruset elimineras alltså av nollställen hos estimatorn. Vidare ingår kanalens nämnarpolynom A alltid i filtrets täljare. Resonanstoppar hos kanalen kompenseras därför av nollställen hos estimatorn.

Spektralfaktorn β kommer att utgöra estimatorns nämnarpolynom, och måste därför vara stabil. Spektralfaktoriseringen (7) utgör en kvadratisk ekvation i koefficienterna hos β . Under Antagande A har ekvationen garanterat en stabil lösning, då dess högerled kan garanteras vara fritt från nollställen på enhetscirkeln.

Ekvationen (7) utgör en beräkning av täljarpolynomet till en innovationsmodell för mätsignalen

$$y(k) = s(k) + w(k) = \frac{BC}{AD}e(k) + \frac{M}{N}v(k) \triangleq \frac{\beta}{DNA}\epsilon(k) \quad (10)$$

⁸Om N och/eller D hade nollställen på enhetscirkeln (t.ex integrerande modeller, "ARIMA-modeller") så skulle inte lösningen ändras, men vi skulle behöva utföra steg 4 i förra avsnittet.

⁹Ekvationen benämns Diofantisk i analogi med heltalsekvationer av typ $ax+by=c$, där a, b, c är kända heltal och x, y är okända heltal. Sådana ekvationer studerades av indiska och grekiska matematiker, bland andra Diofantos, för ca 2500 år sedan. Heltal och polynom lyder likartade räkneregler, då båda algebraiskt sett utgör ringar [11].

Denna modell beskriver mätsignalen som en ARMA-process driven av ett enda ekvivalent vitt brus $\epsilon(k)$. Filtret β/DNA är stabilt och har en stabil invers¹⁰. Utnyttjas innovationsmodellen (10) och antagandet att $D(z^{-1})$ är stabil, så kan estimatorn (9) skrivas som

$$\hat{u}(k|k+m) = \frac{Q_1}{D} \frac{DNA}{\beta} y(k+m) = \frac{Q_1}{D} \epsilon(k+m) \quad (11)$$

Wienerestimatorn funktion kan tolkas som att det först inverterar fram en vit innovationssekvens ur mätsignalen och sedan filtrerar denna sekvens.

Filtret DNA/β brukar kallas för *det vitande filtret*. Det kan verka märkligt att man kan utnyttja en vit sekvens som informationsbärare. Att $\epsilon(k)$ är till nytta för estimering beror på att den må vara okorrelerad med sig själv vid olika tidpunkter, men den är (förhoppningsvis) korrelerad med den önskade signalen $u(k)$! I själva verket kan det klassiska uttrycket för det kausala IIR-Wienerfiltret för detta problem uttryckas på formen

$$\hat{u}(k|k+m) = \left\{ \phi_{u(k)\epsilon(k+m)} \right\}_+ \epsilon(k+m) \quad (12)$$

där $\phi_{u(k)\epsilon(k+m)}$ är kors-spektrum mellan den sökta signalen $u(k)$ och innovationen $\epsilon(k+m)$. Uttrycket $\{ \cdot \}_+$ representerar att endast den *kausala* delen av motsvarande korskovariansfunktion får utnyttjas. Filterblocket $\{ \cdot \}_+$ har således impulssvaret

$$[Eu(k)\epsilon^*(k+m)] + [Eu(k)\epsilon^*(k+m-1)]q^{-1} + [Eu(k)\epsilon^*(k+m-2)]q^{-2} + \dots \quad (13)$$

Jämförs (12) med (11) inses att impulssvaret (13) kan representeras av IIR-filtret

$$\left\{ \phi_{u(k)\epsilon(k+m)} \right\}_+ = \frac{Q_1(q^{-1})}{D(q^{-1})} \quad (14)$$

Lösningen av den Diofantiska ekvationen (8), som ger $Q_1(q^{-1})$, utgör helt enkelt en behändig metod att beräkna den kausala faktorn i Wienerestimatorn.

Man kan visa [1] att det minimala kriterievärdet kan uttryckas som

$$E|\epsilon(k)|_{\min}^2 = \frac{\lambda_e}{2\pi j} \oint_{|z|=1} \left(\left\{ 1 - \mathcal{R}z^m \frac{B}{A} \right\} \frac{CC_*}{DD_*} \left\{ 1 - \frac{B_*}{A_*} z^{-m} \mathcal{R}_* \right\} + \mathcal{R} \frac{M\rho M_*}{NN_*} \mathcal{R}_* \right) \frac{dz}{z} \quad (15)$$

$$= \frac{(\lambda_e/r)}{2\pi j} \oint_{|z|=1} \left(\frac{LL_*}{\beta\beta_*} + \rho \frac{CMAC_*M_*A_*}{\beta\beta_*} \right) \frac{dz}{z} \quad (16)$$

Integranderna representerar estimeringsfelets spektraltäthet. I (15) representerar den första termen bidraget från ofullständig signalmodellering medan den sista termen är mätbrusets bidrag, som försvinner för $\rho = 0$.

¹⁰Om D , N eller A har nollställen på enhetscirkeln så utgör modellen en så kallad generaliserad innovationsmodell [14]. Den är då inte stabil, men har fortfarande stabil invers.

Den optimala estimatorns frekvensegenskaper kommer att bestämmas av signal-till-brusförhållandet i frekvensdomänen. Vid frekvenser ω_1 där bruset har låg effekt prioriteras den första termen i (15), och vi får

$$\mathcal{R}(e^{-j\omega_1})e^{j\omega_1 m} \frac{B(e^{-j\omega_1})}{A(e^{-j\omega_1})} \approx 1$$

om m valts tillräckligt stor.

I uttrycket (16) representerar den första termen estimeringsfel som orsakas av att ett ändligt antal framtida data m utnyttjas. Den sista termen, som återstår då $L_* = 0$, representerar den asymptotiska förlusten då $m \rightarrow \infty$, d.v.s. felet som erhålls med det *ickekausala Wienerfiltret*. Om kausalitetsvillkoret på $\mathcal{R}$ släpps, så kan nämligen ortogonaliteten uppnås genom att integranden i (6) sätts till noll, med $L_* = 0$. Det ickekausala Wienerfiltret erhålls ur (A.8) i Appendix A med $\mathcal{L}_* = 0$, och z utbytt mot q . Om vi förlänger med D erhålls därur

$$\lim_{m \rightarrow \infty} q^m \mathcal{R}(q^{-1}) = \frac{B_* A C C_* N N_*}{r \beta \beta_*} = \left(\frac{B_* C C_* N_*}{r \beta_* D} \right) \left(\frac{D N A}{\beta} \right) . \quad (17)$$

Den sista termen i det sista uttrycket är det vitande filtret, medan den första termen utgör korskovariansfunktionen mellan $u(k)$ och innovationen. Denna funktion har ett dubbelsidigt oändlig impulssvar. En kausal trunkering (för ändligt m) av detta impulssvar ger uttrycken (11)-(14) för det kausala Wienerfiltret.

4 Lösning av syntesekvationerna

Den Diofantiska ekvationen (8) är *garanterat lösbar*, då de kända faktorerna i högerledet saknar gemensamma faktorer då D är stabil (eller har nollställnen på enhetscirkeln). Polynomet $\beta_*(z)$ kommer ju uteslutande att ha rötter *utanför* $|z| = 1$. Att lösningen är *unik* kan visas på följande sätt. Låt $\tilde{Q}_1$ och $\tilde{L}_*$ vara en lösning. Varje lösning till (8) kan då uttryckas som

$$\{Q_1(q^{-1}), L_*(q)\} = \{\tilde{Q}_1(q^{-1}) - qX D(q^{-1}), \tilde{L}_*(q) + X\beta_*(q)\}$$

för något polynom

$$X = x_{-n_1} q^{-n_1} + \dots + x_{-1} q^{-1} + x_0 + x_1 q + \dots + x_{n_2} q^{n_2} ,$$

vilket inses genom direkt insättning i (8). Eftersom filtret måste vara kausalt, så kan $Q_1(q^{-1})$ ej innehålla termer i framåt-skiftoperatorn q . Å andra sidan får $L_*(z)$ p.g.a. ortogonaliteten ej ha några poler i $|z| = 1$, vilket skulle vara fallet om $L_*(q)$ innehöll termer i negativa potenser av q (se Appendixet). Alltså är $X = 0$ det enda alternativet, så lösningen $(\tilde{Q}_1, \tilde{L}_*)$ måste vara unik.

Gradtalen hos de obekanta polynomen bestäms av att de högsta gradtalen i q^{-1} respektive q som förekommer någonstans ska vara "näbara" av variabler. Den högsta graden i q^{-1} i de tre termerna i (8) är

$$m + nc \quad , \quad \deg Q_1 \quad , \quad nd - 1$$

där nc är graden hos $C(q^{-1})$ osv. Den högsta graden i positiva potenser av q i de tre termerna är

$$nc + nb + nn - m \quad , \quad n\beta, \quad \deg L_* + 1$$

Detta ger gradtalsvillkoren på de obekanta polynomen

$$\deg Q_1 = \max(nc + m, nd - 1) \quad ; \quad \deg L_* = \max(nc + nb + nn - m, n\beta) - 1 \quad . \quad (18)$$

Ekvationen (8) kan då omformas till ett linjärt ekvationssystem med lika många ekvationer som obekanta.

Man kan lägga märke till att för prediktorer (negativa m) så kommer filtrets komplexitet (graden hos Q_1) ej att påverkas av prediktionshorisontens längd. För smoothers ökar estimatorns komplexitet med m .

Spektralfaktoriseringar har analytisk lösning endast för polynom av grad högst 2. Om högerledet till (7) ges av

$$g_o + g_1(q + q^{-1}) + g_2(q^2 + q^{-2})$$

där koefficienterna är reella, så ges skalären r och polynomet $\beta(q^{-1}) = 1 + \beta_1 q^{-1} + \beta_2 q^{-2}$ av [13]

$$\gamma \triangleq \frac{g_o}{2} - g_2 + \sqrt{\left(\frac{g_o}{2} + g_2\right)^2 - g_1^2}$$

$$r = \left(\gamma + \sqrt{\gamma^2 - 4g_2^2}\right) / 2 \quad ; \quad \beta_1 = \frac{g_1}{r + g_2} \quad ; \quad \beta_2 = \frac{g_2}{r} \quad . \quad (19)$$

För högre gradtal måste numeriska lösningar tillgripas. Några alternativ är:

- Man kan beräkna högerledets rötter och bilda $\beta(q^{-1})$ ur faktorerna med rötter innanför enhetscirkeln. Denna metod är implementerad i `spefac2.m`. Högerledet i (7) har nollställen symmetriskt fördelade m.a.p enhetscirkeln.
- Man kan lösa ekvationen iterativt, med en Newton-Raphsonmetod som har kvadratisk konvergens [9, 10]. Algoritmen är implementerad i `spefac.m`.
- Ekvationerna kan lösas indirekt via en algebraisk Riccatiekvation [12]. Algoritmen har implementerats i `spefac3.m`.

Figur 3 visar anropet till en matlabfunktion `fnfilt.m`, som kan lösa det ovan skisserade Wienerproblemet, och som även kan hantera filtrering av den önskade signalen.

5 Generaliseringar

Den ovan skisserade lösningen kan generaliseras till *flervariabla system*, Metodiken kan även generaliseras till fall då *robusta filter* måste konstrueras för mängder av tänkbara signalmodeller, istället för en enda unik modell [15, 17]. Signalmodeller är ju aldrig exakt kända, och ofta kan modellosäkerheten vara stor. Ett möjligt

```

function [Q0,R0] = fnfilt(C0,D0,B0,A0,M0,N0,S,T,rho,m,k)

Computes NOMINAL WIENER FILTER Q0/R0 from a nominal signal model

          B0          M0          C0
y(t) = ----u(t-k) + ----v(t) ;   u(t) = ----e(t)   ;
          A0          N0          D0

E|v|^2/E|e|^2 = rho

Estimated signal:   f(t-m)= (S/T)u(t-m)
Estimator:         fhat(t-m|t) = (Q0/R0)y(t).

The polynomials may have complex coefficients.
USES: abstar, addcenter, spefac2, polysolve.
Author: Mikael Sternad

```

Figure 3: Matlab-funktion fnfilt.m för Wienerfiltrering, prediktion och smoothing av skalära signaler. Observera att kanalens tidsfördröjning kan anges separat.

angreppssätt är i sådana fall att försöka optimera prestanda för det värsta fallet. En annan lösningsmetodik går ut på att medelvärdesbilda kriteriet för olika signalmodeller och minimera detta medelvärde.

Slutligen kan man använda liknande metoder för att konstruera filter som minimerar den *högsta toppen i felsignalens spektrum*, istället för felsignalens varians. Sådana estimatorer kallas H_∞ -filter [7].

En begränsning hos Wiener-metodiken är att det är ganska besvärligt att generalisera spektralfaktoriseringsekvationer till tidsvariabla polynom. För tidsvariabla system tillgriper man hellre Kalmanfiltrering.

En annan begränsning som Wiener- och Kalmanfilter har gemensamt är att estimatorerna är linjära. Om signal och brus är normalfördelade kan man dock visa att de utgör de bästa estimatorerna, även bland alla olinjära estimatorer, om respektive signalmodeller är korrekta. Om signal och/eller brus är icke-gaussiska så utgör de dock fortfarande de bästa *linjära* estimatorerna. Det kan dock finnas olinjära estimatorer som är överlägsna. Ett exempel på detta är digitala kommunikationssystem, där $u(k)$ är en symbol med icke-gaussisk fördelning. Ett annat exempel är om enstaka stora brusvärden kan förekomma, dvs om brusets fördelningsfunktion har långa "svansar". Man kan då förbättra prestanda genom att låta en statisk olinjär funktion verka på innovationssignalen i Wienerfiltret och klippa av dess högsta toppar, som troligen orsakas av brus.

6 Ett numeriskt exempel

Exemplet är hämtat från Kap III.I i [3]. Vi använder en modell med

$$C(q^{-1}) = A(q^{-1}) = M(q^{-1}) = N(q^{-1}) = 1$$

$$D(q^{-1}) = 1 - 0.5q^{-1} \quad ; \quad B(q^{-1}) = 1 - 1.4q^{-1} + 0.92q^{-2} \quad .$$

Brusvarianserna är $\lambda_e = 1$, $\lambda_v = \rho = 0.01$ och ett filter ska beräknas, dvs $m = 0$. Vi ska alltså bygga ett utjämningsfilter för kanalen $B(q^{-1})$, för att skatta signalen $u(k)$ som har autoregressiv statistik. Spektralfaktoriseringen blir

$$r\beta\beta_* = BB_* + \rho AA_* = (1 - 1.4q^{-1} + 0.92q^{-2})(1 - 1.4q + 0.92q^2) + 0.01(1 - 0.5q^{-1})(1 - 0.5q) \quad .$$

I detta fall blir spektralfaktoriseringen av ordning 2, så en analytisk lösning existerar. Lösningen ges av (19), där

$$g_o + g_1(q + q^{-1}) + g_2(q^2 + q^{-2}) = 3.8189 - 2.6930(q + q^{-1}) + 0.92(q^2 + q^{-2})$$

ger $\gamma = 1.8575$ och

$$r = 1.0560 \quad ; \quad \beta(q^{-1}) = 1 - 1.3629q^{-1} + 0.8712q^{-2} \quad . \quad (20)$$

Uttrycken (18) ger att $Q_1(q^{-1})$ ska ha grad $nd - 1 = 0$ och $L_*(q)$ får grad $n\beta - 1 = 1$. Vi ska alltså lösa den Diofantiska ekvationen

$$1 - 1.4q + 0.92q^2 = 1.056(1 - 1.3629q + 0.8712q^2)Q_{10} + (-0.5 + q)(\ell_o + \ell_1q) \quad . \quad (21)$$

Genom att sätta termer med samma grad i q lika i höger och vänster led erhåller vi ett ekvationssystem med block-Toeplitz-struktur:

$$\begin{pmatrix} 1 \\ -1.4 \\ 0.92 \end{pmatrix} = \begin{pmatrix} 1.056 & -0.5 & 0 \\ -1.4392 & 1 & -0.5 \\ 0.92 & 0 & 1 \end{pmatrix} \begin{pmatrix} Q_{10} \\ \ell_o \\ \ell_1 \end{pmatrix} \quad . \quad (22)$$

Lösningen är

$$\begin{aligned} Q_1(q^{-1}) = Q_{10} &= 0.9357 \\ L_*(q) &= -0.0238 + 0.0592q \quad . \end{aligned}$$

Det optimala filtret (11) blir alltså ett andra ordningens IIR-filter

$$\hat{u}(k|k) = \frac{Q_1(q^{-1})}{\beta(q^{-1})} y(k) = \frac{0.9357}{1 - 1.3629q^{-1} + 0.8712q^{-2}} y(k) \quad (23)$$

Filtrets överföringsfunktion är nästan en invers av kanalen $s(k) = B(q^{-1})u(k)$. Estimeringsfelets varians blir, med $C = M = A = 1$, enligt (16)

$$J_{\min} = E|\varepsilon(k)|_{\min}^2 = \frac{(\lambda_e/r)}{2\pi j} \oint_{|z|=1} \frac{LL_* + \rho}{\beta\beta_*} \frac{dz}{z} \quad . \quad (24)$$

Med $\rho = 0.01$ och $L_*(q)$ enligt ovan, blir $J_{\min} = 0.1006$.

A Härledning av Wienerestimatoren

1. Härled en innovationsmodell för mätsignalen $y(k)$:

$$y(k) = s(k) + w(k) = \frac{BC}{AD}e(k) + \frac{M}{N}v(k) = \frac{\beta}{ADN}\epsilon(k) \quad (\text{A.1})$$

där $\epsilon(k)$ är en vit sekvens med medelvärde noll och varians λ_ϵ (*innovationen*) och där $\beta(q^{-1})$ är ett moniskt och stabilt polynom som ska bestämmas. Spektra hos alla led i (A.1) måste vara lika, vilket ger

$$\begin{aligned} & \frac{B(z^{-1})C(z^{-1})}{A(z^{-1})D(z^{-1})}\lambda_\epsilon \frac{B_*(z)C_*(z)}{A_*(z)D_*(z)} + \frac{M(z^{-1})}{N(z^{-1})}\lambda_v \frac{M_*(z)}{N_*(z)} \\ &= \frac{\beta(z^{-1})\lambda_\epsilon\beta_*(z)}{A(z^{-1})A_*(z)D(z^{-1})D_*(z)N(z^{-1})N_*(z)} . \end{aligned}$$

Om en skalär r och en bruskvot ρ införs via

$$r = \lambda_\epsilon/\lambda_e \quad ; \quad \rho = \lambda_v/\lambda_e \quad (\text{A.2})$$

och vi gör liknämngt så erhålls *spektralfaktoriseringen* (7)

$$r\beta\beta_* = CC_*BB_*NN_* + \rho MM_*DD_* . \quad (\text{A.3})$$

2. Inför en variationsterm och utnyttja Parsevals formel.

Med estimeringsfelet (4) får vi

$$\varepsilon(k) = \left(1 - \mathcal{R}q^m \frac{B}{A}\right) \frac{C}{D}e(k) - q^m \mathcal{R} \frac{M}{N}v(k) . \quad (\text{A.4})$$

Variationstermen som adderas till estimatet blir

$$\nu(k) = \mathcal{T}y(k+m) = \mathcal{T}q^m \left(\frac{BC}{D}e(k) + \frac{M}{N}v(k) \right) . \quad (\text{A.5})$$

Notera att $e(k)$ och $v(k)$ antas okorrelerade. Uttrycket (6) blir då

$$\begin{aligned} & E\varepsilon(k)\nu(k)^* = \\ & E \left(\left\{1 - \mathcal{R}q^m \frac{B}{A}\right\} \frac{C}{D}e(k) \right) \left(\mathcal{T}q^m \frac{BC}{AD}e(k) \right)^* - q^m \mathcal{R} \frac{M}{N}v(k) \left(\mathcal{T}q^m \frac{M}{N}v(k) \right)^* \\ &= \frac{1}{2\pi j} \oint_{|z|=1} \frac{[\{A - z^m \mathcal{R}B\} z^{-m} C \lambda_\epsilon C_* B_* N N_* - \mathcal{R} M \lambda_v M_* A A_* D D_*]}{D D_* A A_* N N_*} \mathcal{T}_* \frac{dz}{z} \\ &= \frac{\lambda_e}{2\pi j} \oint_{|z|=1} \frac{[z^{-m} A C C_* B_* N N_* - \mathcal{R} r \beta \beta_*]}{A A_* D D_* N N_*} \mathcal{T}_* \frac{dz}{z} \quad (\text{A.6}) \end{aligned}$$

där vi utnyttjade (A.2) och (A.3) i sista ledet.

3. Säkerställ ortogonaliteten genom att eliminera integrandens alla poler innanför enhetscirkeln.

I nämnaren i (A.6) kommer nollställena till $D_*(z)$, $A_*(z)$ och $N_*(z)$ alla att ligga *utanför* enhetscirkeln, eftersom $D(z^{-1})$, $A(z^{-1})$ och $N(z^{-1})$ har alla nollställena *inuti* enhetscirkeln enligt antagande A. Polerna hos den rationella funktionen $\mathcal{T}_*(z)$ ligger alla utanför enhetscirkeln, eftersom variationsfiltret $\mathcal{T}(z^{-1})$ måste antas vara stabilt.

Faktorn $A(z^{-1})D(z^{-1})N(z^{-1})$ i nämnaren i (A.6) samt polen i origo från nämnarfaktorn z måste dock elimineras. Dessutom får vi inte ha kvar några polynom i z^{-1} , bara polynom i z .¹¹ Skriv nu integranden i (A.6) som

$$\left(\frac{z^{-m} ACC_* B_* N N_* - \mathcal{R} r \beta \beta_*}{ADNz} \right) \frac{1}{A_* D_* N_*} \mathcal{T}_* . \quad (\text{A.7})$$

Vi säkerställer $E\varepsilon(k)\nu(k)^* = 0$ om (och endast om) faktorn i parentes utgör ett polynom eller rationell funktion $\mathcal{L}_*(z)$ med poler endast utanför enhetscirkeln, dvs om

$$z^{-m} ACC_* B_* N N_* - \mathcal{R} r \beta \beta_* = ADNz \mathcal{L}_*(z) . \quad (\text{A.8})$$

Polynomen A och N måste ingå i täljaren hos $\mathcal{R}$, då de ingår i de andra termerna i (A.8). Nämnaren hos $\mathcal{R}$ kan endast innehålla faktorn β . Andra nämnarfaktorer skulle behöva ingå som nämnarfaktorer i alla övriga termer i (A.8). Ingen annan term i (A.8) innehåller dock några rationella funktioner med poler inuti enhetscirkeln. Alltså gäller

$$\mathcal{R}(z^{-1}) = \frac{Q_1(z^{-1})N(z^{-1})A(z^{-1})}{\beta(z^{-1})} \quad (\text{A.9})$$

där polynomet $Q_1(z^{-1})$ återstår att bestämma. Ekvationen (A.8) förenklas nu till

$$z^{-m} CC_* B_* N_* = r \beta_* Q_1 + z D z \mathcal{L}_* . \quad (\text{A.10})$$

Då ingen av de andra termerna innehåller nämnare, måste funktionen $\mathcal{L}_*(z)$ vara ett polynom i z , $L_*(z)$. Vi får då designekvationen (8) om z byts mot q .

¹¹Polynom i z^{-1} svarar nämligen mot rationella funktioner i z med ett antal poler i origo. Till exempel gäller

$$1 + 0.5z^{-1} = \frac{z + 0.5}{z} .$$

References

- [1] A. Ahlén and M. Sternad, “Optimal deconvolution based on polynomial methods,” *IEEE Transactions on Signal Processing*, vol. 37, pp. 217–226, 1989.
- [2] A. Ahlén and M. Sternad, “Wiener filter design using polynomial equations,” *IEEE Transactions on Signal Processing*, vol. 39, pp. 2387–2399, 1991.
- [3] A. Ahlén and M. Sternad, “Derivation and design of Wiener filters using polynomial equations,” In C T Leondes ed., *Control and Dynamic Systems*, vol 64: Stochastic Techniques in Digital Signal Processing, Academic Press, New York, 1994, pp 353–418.
- [4] B.D.O. Anderson and J. B. Moore, *Optimal Filtering*. Prentice Hall, Englewood Cliffs, NJ, 1979.
- [5] H.W. Bode and C.E. Shannon, “A simplified derivation of linear least square smoothing and prediction theory,” *Proceedings of the I.R.E.*, vol. 38, pp. 417–425 (1950).
- [6] M.J. Grimble, “Polynomial systems approach to optimal linear filtering and prediction,” *International Journal of Control*, vol. 41, pp. 1545–1564, 1985.
- [7] M.J. Grimble and A. ElSayed, “Solution of the H_∞ optimal linear filtering problem for discrete-time systems,” *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, pp. 1092–1104 (1990).
- [8] M.H. Hayes, *Statistical Digital Signal Processing and Modeling*. Wiley, New York, 1996.
- [9] J. Ježek and V. Kůčera, “Efficient algorithm for matrix spectral factorization,” *Automatica*, vol. 21, pp. 663–669, 1985.
- [10] V. Kůčera, *Discrete Linear Control*. Wiley, Chichester, 1979.
- [11] V. Kůčera, *Analysis and Design of Discrete Linear Control Systems*. Academia, Prague, 1991.
- [12] B.D.O. Andersson, “An algebraic solution to the spectral factorization problem,” *IEEE Transactions on Automatic Control*, vol. 4, pp. 410–414, 1967.
- [13] V. Peterka, “Predictor-based self-tuning control”. *Automatica*, vol 20, pp 39–50 (1984).
- [14] U. Shaked, “A generalized transfer function approach to linear stationary filtering and steady-state optimal control problems,” *International Journal of Control*, vol 29, pp. 741–770, 1976.
- [15] M. Sternad and A. Ahlén, “Robust filtering and feedforward control based on probabilistic descriptions of model errors,” *Automatica*, vol 29, no 3, pp. 661–679, 1993.
- [16] N. Wiener, *Extrapolation, Interpolation and Smoothing of Stationary Time Series*. The Technology Press and Wiley, New York, 1950.
- [17] K. Öhrn, A. Ahlén and M. Sternad, “A probabilistic approach to multivariable robust filtering and open-loop control.” *IEEE Transactions on Automatic Control*, vol. 40, pp. 405–417 , 1995.