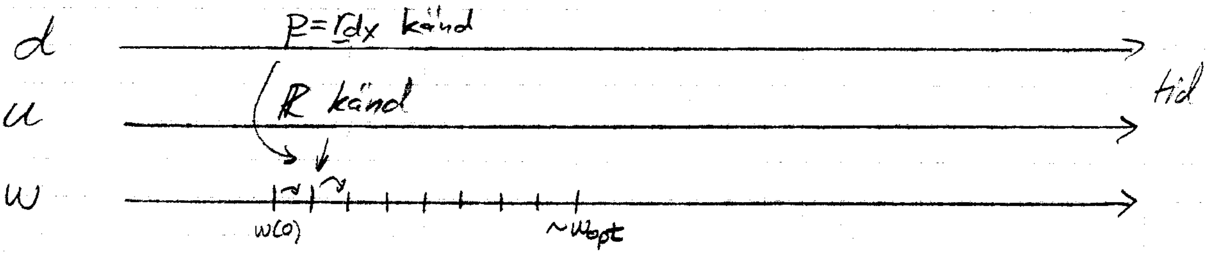
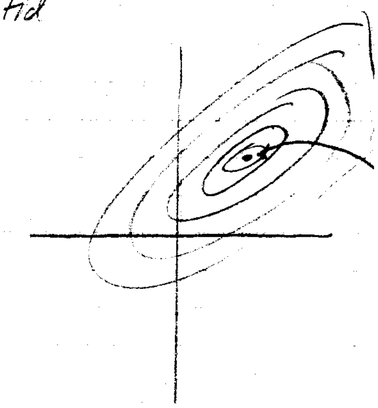
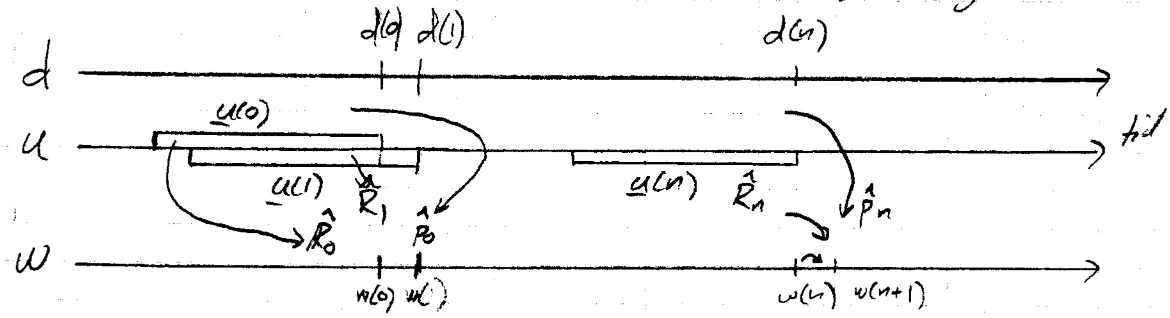
Steepest descent metoden (WSS → R finns + känd)





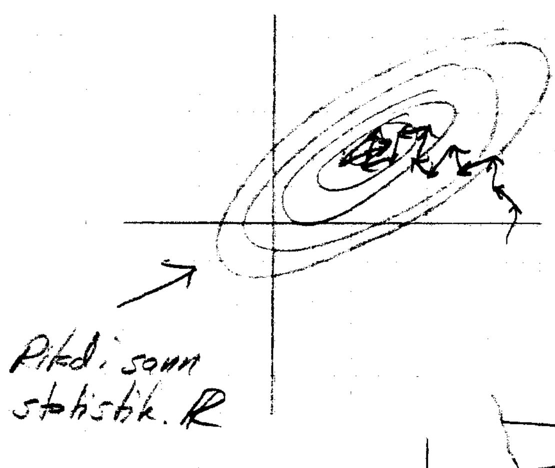
LMS (adaptivt filter) (WSS ⟹ R finns men okänd)
Måste skattas från signaler

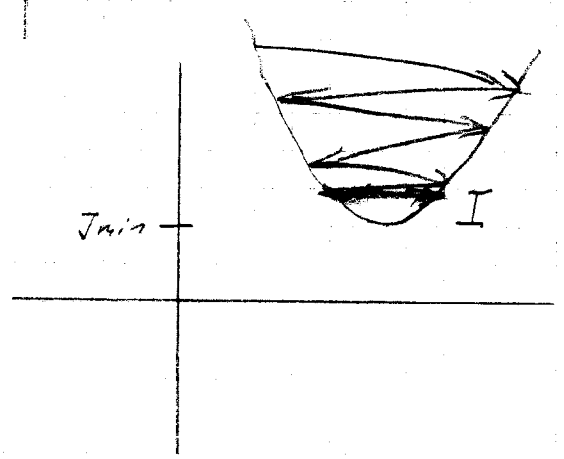


$$\hat{R}_0 = u(0)u^H(0)$$
$$\hat{R}_1 = u(1)u^H(1)$$

$$R_i \neq R_j$$

- Även om statistiken i medel ges av R och p så behöver inte $\hat{R}_i = R$ och $\hat{p}_i = p$
- Varje ny insignalblock har en egen uppfattning om var skålens botten är.
- Eftersom det finns en medelstatistik så är skålarna hyfsat lika
- Kan anpassa sig till förändrad statistik
- Innan iterations index, men för LMS tidsindex.

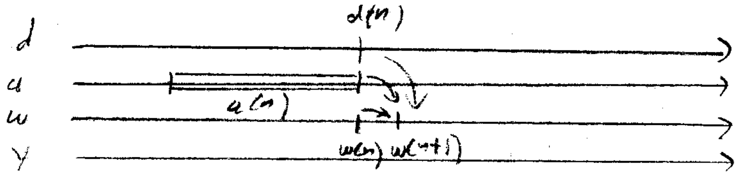




Bruset i $\hat{R}_i$ och $\hat{p}_i$ gör att metoden aldrig riktigt når botten.

Verktyg

Independence theory:



Antaganden

1. Datavektorerna statistiskt oberoende (ny info i varje $u(n)$)

$$u(n) \perp u(m) \quad n \neq m$$

2. Datavektorn vid tidpunkt n stat. ober. av önskade signalen fram till $n-1$

$$u(n) \perp d(m) \quad m = 0..n-1$$

3. Den önskade signalen vid n beror av datavektorn fram till n men inte av den önskade signalen fram till $n-1$

$$d(n) \perp d(m) \quad m = 0..n-1$$

Konsekvenser: 1. $\hat{w}(n) = \hat{w}(n-1) + \mu u(n-1)\left(d(n-1) - w^H(n-1)u(n-1)\right)$

$$\Rightarrow \hat{w}(n) \perp u(n) \qquad (I_a)$$

2. $e_o(n) = d(n) - w_o^H u(n)$ $\quad$ $e_o(n)$ är den del av $d(n)$ som inte kan repr. av $u(n)$

$$\Rightarrow e_o(n) \perp u(n) \quad (I_b)$$

$$e_o(n) \perp \hat{w}(n) \quad (I_c)$$ gäller även $\varepsilon(n)$

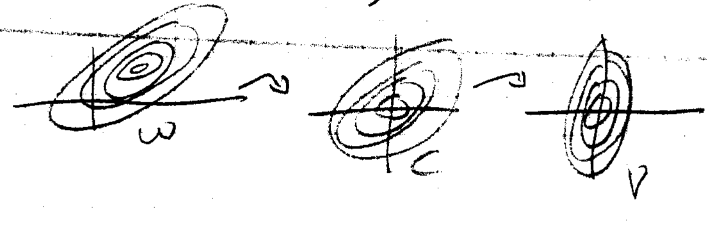
$\tilde{d}(n)$ $\quad$ $\tilde{u}(m), d(m)$ $\quad m = 0..n-1$

Matristeori

$$tr(AB) = tr(BA) \quad (M_a)$$

$$E[trA] = tr(E[A]) \quad (M_b)$$

Konvergensanalys SD (WSS)



1 Introducera $c(n) = w(n) - w_o$

$$v(n) = Q^H(w(n) - w_o)$$

2 Uttryck uppdateringen i $v(n)$

$$w(n+1) = w(n) + \mu(p - Rw(n))$$

$$\Rightarrow v(n) = (I - \mu\Lambda)v(n)$$ — $\left\{ \begin{array}{l} \text{oberoende rader} \\ \text{homogen} \\ \text{kopplad till } \lambda_i \end{array} \right.$

forts.

2 forts. Eftersom raderna är oberoende så kan insvängningar för de olika egenmoderna separeras:

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

Lösning

$$v_k(n) = (\underbrace{1-\mu\lambda_k}_{<1})^n v_k(0)$$

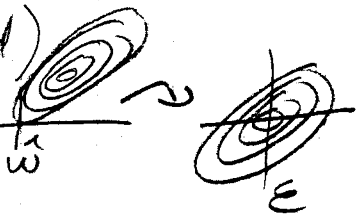
stabilitet

$$0 < \mu < \frac{2}{\lambda_{max}}$$

3. Uttryck $J(n)$ i $v(n)$ för att se hur medelkvadratfelet konvergerar mot noll.

$$J(n) = J_{min} + \sum_k \lambda_k |v_k(n)|^2$$
$$= J_{min} + \sum_k \lambda_k (1-\mu\lambda_k)^{2n} |v_k(0)|^2$$

Konvergensanalys LMS (WSS, R finns men är okänd)



1. Introducera koefficientvektorn $\varepsilon(n) = \hat{w}(n) - w_0$

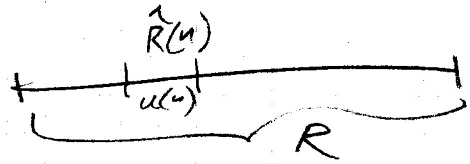
2. Uttryck uppdatering i $\varepsilon(n)$

$$\hat{w}(n+1) = \hat{w}(n) + \mu u(n) e^*(n)$$

$$\varepsilon(n+1) = \varepsilon(n) + \mu u(n)\left(d^*(n) - u^H(n)\varepsilon(n) - u^H(n) w_0\right)$$

$$\varepsilon(n+1) = \left(I - \mu u(n) u^H(n)\right)\varepsilon(n) + \mu u(n) e_o^*(n)$$

- ej ober rader
- ej homogen
- ej naturligt kopplad till λ_i



Direct Averaging: Ersätt $\hat{R}$ med R. Om μ litet kommer dessa båda diff. ekv att bete sig lika.

$$\Rightarrow \varepsilon(n+1) = (I - \mu R)\varepsilon(n) + \mu u(n) e_o^*(n) \qquad *$$

- kopplad till λ_i genom R men också till $u(n)$

3. Uttryck $J(n)$ i $\varepsilon(n)$

$$J(n) = E\left[|e(n)|^2\right] = E\left[|d(n) - \hat{w}^H(n) u(n)|^2\right]$$
$$= E\left[|d(n) - \varepsilon^H(n) u(n) - w_o^H u(n)|^2\right] =$$
$$= E\left[|e_o(n) - \varepsilon^H(n) u(n)|^2\right] = [I_b, I_c]$$
$$= E\left[|e_o(n)|^2\right] + E\Big[\underbrace{\varepsilon^H(n) u(n) u^H(n) \varepsilon(n)}_{\text{skalär}}\Big]$$
$$= [M_a, M_b] = E\left[|e_o(n)|^2\right] + tr\left(E\left[u(n) u^H(n) \varepsilon(n) \varepsilon^H(n)\right]\right)$$
$$= [I_a] = E\left[|e_o(n)|^2\right] + tr\Big(\underbrace{E\left[u(n) u^H(n)\right]}_{R} E\left[\varepsilon(n) \varepsilon^H(n)\right]\Big)$$
$$= J_{min} + tr(R K(n)) \quad \text{där} \quad K(n) = E\left[\varepsilon(n)\varepsilon^H(n)\right]$$

4. Diff ekv i $K(n)$

$$K(n+1) = E\left[\varepsilon(n+1)\varepsilon^H(n+1)\right] \overset{*}{=} \left[I_a, I_c, K(n) = E\left[\varepsilon(n)\varepsilon^H(n)\right]\right]$$

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

- matrisform
- rad och kolonnberoende
- inhomogen
- naturligt kopplad till λ_i

5. Separera egenmoder $\qquad$ $K(n)$ korr.matris för filterfelvektorn

$$Q^H R Q = \Lambda$$

$$Q^H K(n) Q = X(n) \qquad \text{(applicera } Q \text{ på } K(n))$$

$$J(n) = J_{min} + tr(RK(n))$$
$$= J_{min} + tr(Q\Lambda Q^H Q X(n) Q^H$$
$$= J_{min} + tr(\Lambda X(n)) \qquad X(n) \text{ likformig med } K(n)$$
$$= J_{min} + \sum_i \lambda_i x_i(n)$$

där $x_i(n)$ = diagonalelement i den likformiga korrmatrisen för filterfelsvektorn

Diff.ekv i $X(n)$

$$X(n+1) = Q^H K(n+1) Q = Q^H(I-\mu R)QX(n)Q^H(I-\mu R)Q + \mu^2 Q^H R Q J_{min}$$
$$= (I-\mu\Lambda)X(n)(I-\mu\Lambda) + \mu^2 \Lambda J_{min}$$

- inhomogen
- oberoende rader

För varje diagonalelement gäller ekv

$$x_i(n+1) = (1-\mu\lambda_i)^2 x_i(n) + \mu^2 \lambda_i J_{min}$$

Slutsats

$$J(n) = J_{min} + \overbrace{\sum_i \lambda_i x_i(n)}^{J_{ex}(n)}$$

där $x_i(n)$ är lösningen till

$$x_i(n+1) = (1-\mu\lambda_i)^2 x_i(n) + \underbrace{\mu^2 \lambda_i J_{min}}_{>0}$$

Observationer

1/ Konvergerar om $|1-\mu\lambda_i| < 1$
dvs $0 < \mu < \frac{2}{\lambda_{max}}$

2 Inhomogen: Konv inte mot noll

$$J(n) = J_{min} + \underbrace{\overset{\to 0}{J_{transient}(n)}} + \underbrace{J_{ex}(\infty)}_{\sim \mu, \lambda_i, J_{min}}$$

Skatta egenvärdena med hjälp av totala energin i insignalvektorn

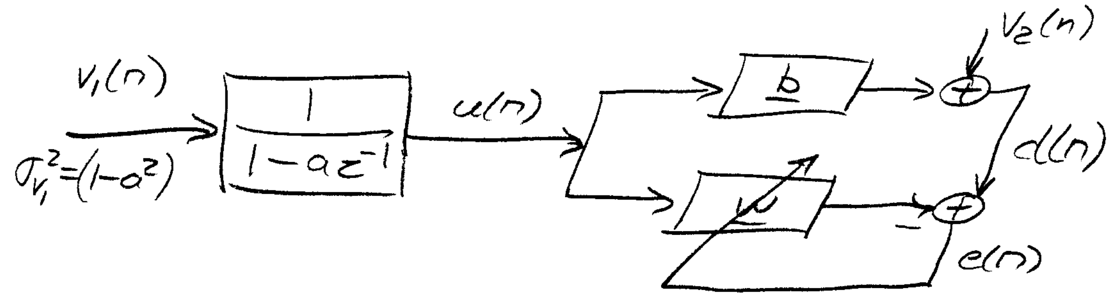
$$\underline{u}^H(n)\underline{u}(n) = \sum_{k=0}^{M-1} |u(n-k)|^2 \approx \sum_{k=0}^{M-1} \underbrace{E[|u(n-k)|^2]}_{r(0)} = M r(0)$$

$$= tr(R) = tr(Q \Lambda Q^H) = tr \Lambda = \sum_{k=0}^{M-1} \lambda_k$$

Byt $0 < \mu < \frac{2}{\lambda_{max}}$ (som gäller i medel)

mot $0 < \mu < \frac{2}{\sum_{k=0}^{M-1} \lambda_k}$ (snävare gräns)

Exempel: Egenvärdesspridning



$$w_0 = \mathbb{R}_u^{-1}\mathbb{p}$$

$$P_u(z) = (1-a^2)\left(\frac{1}{1-az^{-1}}\right)\left(\frac{1}{1-az}\right) \Leftrightarrow r_u(k) = a^{|k|} \quad (r_u(0)=1)$$

$$\Rightarrow \mathbb{R}_u = \begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix}$$

$$p(-k) = E[u(n-k)d(n)] = E[u(n-k)(u(n) * b(n) + v_2(n))]$$

$$= E[u(n-k)(b_0u(n) + b_1u(n-1) + v_2(n))$$

$$= b_0 r_u(-k) + b_1 r_u(-k+1) \qquad k=0..1$$

$$\Rightarrow \mathbb{p} = \begin{bmatrix} b_0 + b_1a \\ b_0a + b_1 \end{bmatrix} \qquad w_0 = \mathbb{R}^{-1}\mathbb{p} = \frac{1}{1-a^2}\begin{bmatrix} 1 & -a \\ -a & 1 \end{bmatrix}\begin{bmatrix} b_0+b_1a \\ b_0a+b_1 \end{bmatrix} = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

$$J_{min} = E[\underbrace{(d(n) - \hat{d}(n))}_{e(n)}d(n)] = E[d(n)d(n)] - E[d(n)\hat{d}(n)]$$

$$= E[(b_0u(n) + b_1u(n-1) + v_2(n))(b_0u(n) + b_1u(n-1) + v_2(n))]$$

$$- E[(b_0u(n) + b_1u(n-1) + v_2(n))(b_0u(n) + b_1u(n-1))] = r_{v_2}(0) = \sigma_{v_2}^2$$

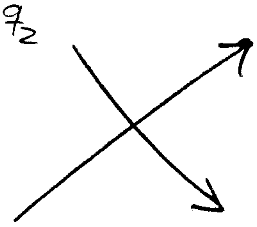
$$|\lambda I - \mathbb{R}_u| = 0 \Rightarrow (\lambda-1)^2 - a = 0 \Rightarrow \lambda = 1 \pm |a|$$

$$\chi(\mathbb{R}_u) = \frac{\lambda_{max}}{\lambda_{min}} = \frac{1+|a|}{1-|a|}$$

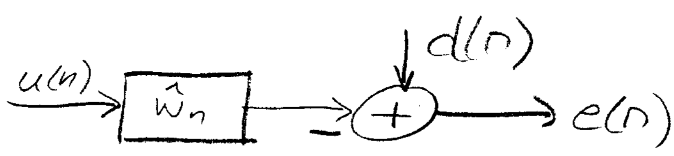
$$\lambda = 1+|a|: \quad \lambda q = \mathbb{R}_u q \Rightarrow (1+|a|)\begin{bmatrix} q_1 \\ q_2 \end{bmatrix} = \begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix}\begin{bmatrix} q_1 \\ q_2 \end{bmatrix}$$

$$\begin{cases} (1+|a|)q_1 = q_1 + aq_2 \\ (1+|a|)q_2 = aq_1 + q_2 \end{cases} \Rightarrow q_1 = \frac{a}{|a|}q_2$$

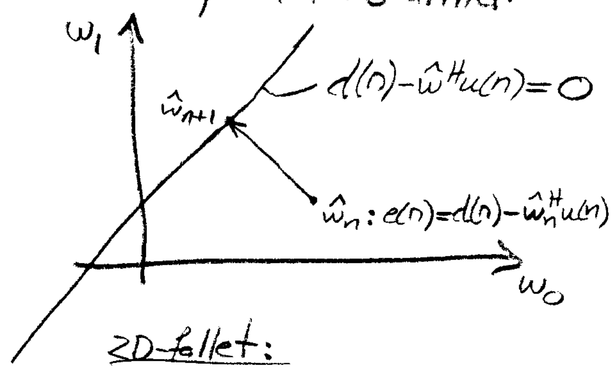
$$\lambda = (1-|a|): \quad q_1 = -\frac{a}{|a|}q_2 \qquad Q = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$



Normalized LMS (tolkning)



För varje n finns alltid:



Strategi: Uppdatera filterkoef så att vi uppfyller $d(n)-\hat{w}_{n+1}^H u(n)=0$ genom att hoppa ett så kort steg som möjligt.

Nackdel: Ingen medelvärdesbildning. Ej robust.

2D-fallet:

$$d(n)-w_0 u(n)-w_1 u(n-1)=0$$

$$w_1=\frac{d(n)}{u(n-1)}-w_0\frac{u(n)}{u(n-1)}$$

$$J(n)=\|\hat{w}_{n+1}-w_n\|^2+\lambda\left(d(n)-\hat{w}_{n+1}^H u(n)\right)+\lambda^*\left(d^*(n)-u^H(n)w_{n+1}^H\right)$$

$$=\hat{w}_{n+1}^H\hat{w}_{n+1}-\hat{w}_{n+1}^H\hat{w}_n-\hat{w}_n^H\hat{w}_{n+1}+\hat{w}_n^H\hat{w}_n+\lambda\left(d(n)-\hat{w}_{n+1}^H u(n)\right)+\lambda^*\left(d^*(n)-u^H(n)\hat{w}_{n+1}\right)$$

$$\begin{cases}\dfrac{\partial J(n)}{\partial \hat{w}_{n+1}^H}=\left(\hat{w}_{n+1}-\hat{w}_n\right)-\lambda u(n)=0 & (1)\\[2ex] \dfrac{\partial J(n)}{\partial \lambda^*}=d^*(n)-u^H(n)\hat{w}_{n+1}=0 & (2)\end{cases}$$

(1) Mult $u^H(n)$ fr. vä.

$$\lambda=\frac{u^H(n)\left(\hat{w}_{n+1}-\hat{w}_n\right)}{\|u(n)\|^2}\overset{(2)}{=}\frac{d^*(n)-u^H(n)\hat{w}_n}{\|u(n)\|^2}=\frac{e^*(n)}{\|u(n)\|^2}$$

(1) ger

$$\hat{w}_{n+1}=\hat{w}_n+\frac{1}{\|u(n)\|^2}\cdot u(n)e^*(n)$$

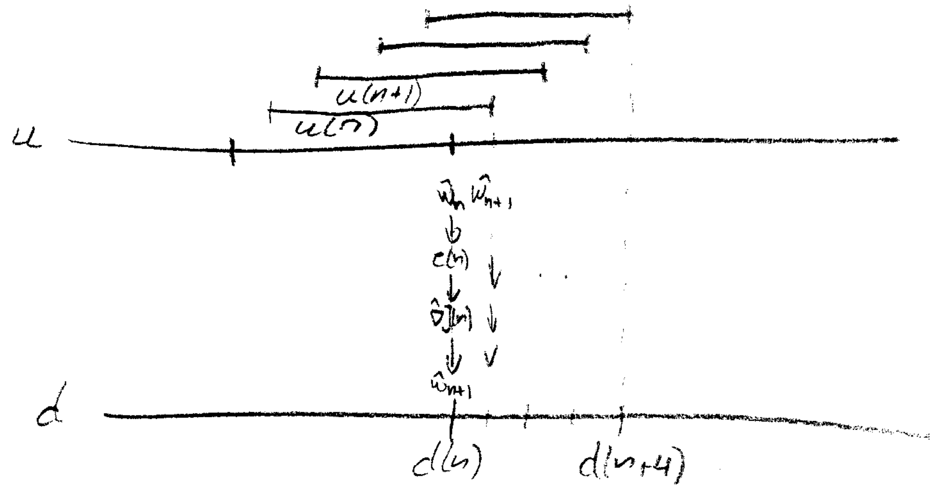
Fix 1: $\tilde{\mu}$ steglängd. Ett kortare steg ökar medelvärdesbildning

Fix 2: a om $\|u(n)\|\to 0$

$\Rightarrow$

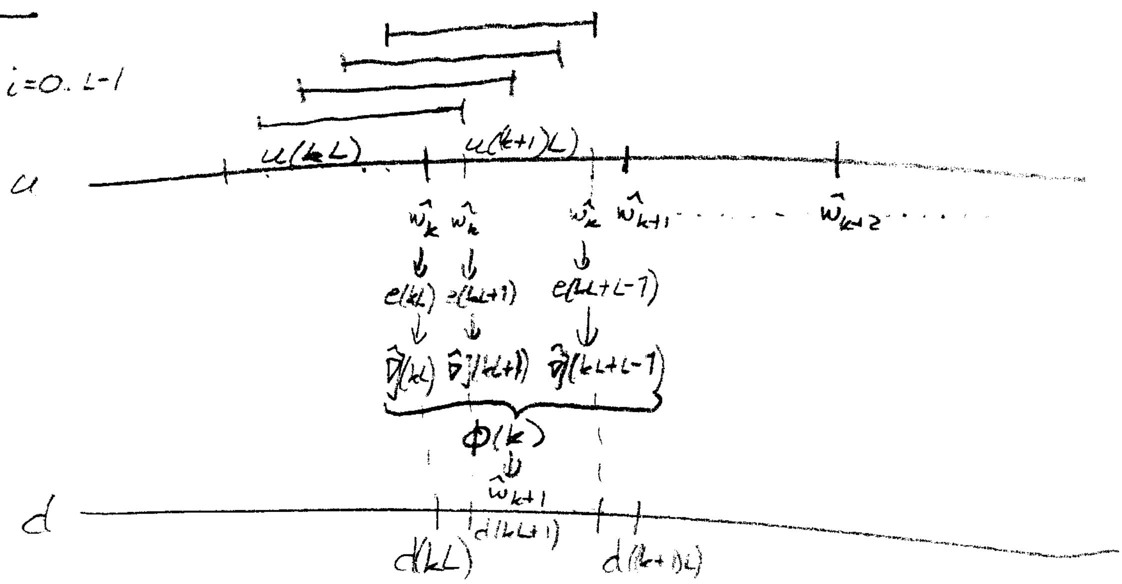
$$\hat{w}_{n+1}=\hat{w}_n+\frac{\tilde{\mu}}{a+\|u(n)\|^2}u(n)e^*(n)$$

LMS

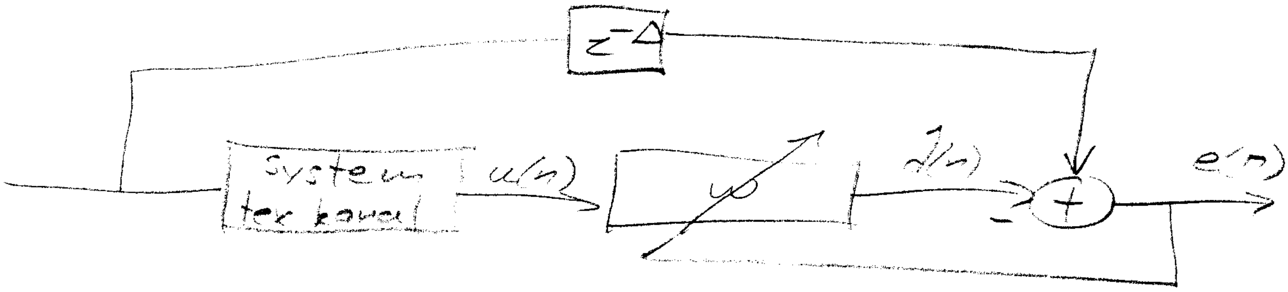


Block LMS

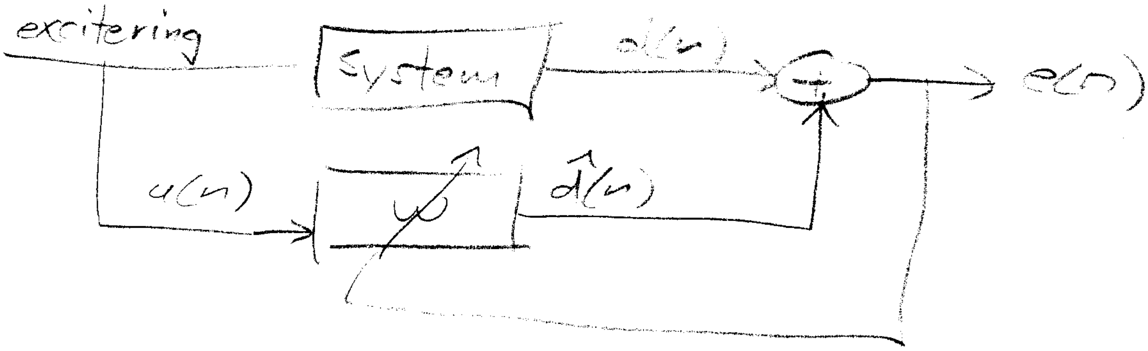
$n = kL + i,\ i = 0..L-1$



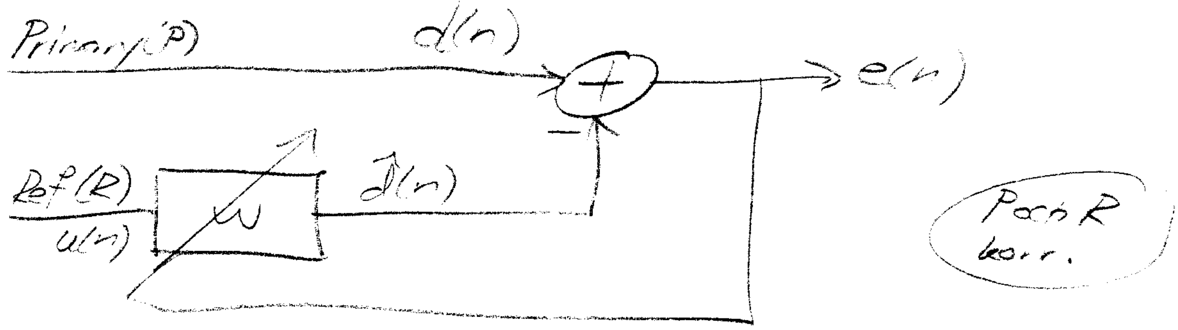
Invers adaptiv modellering, utjämning, avfaltning



Adaptiv modellering, systemidentifiering



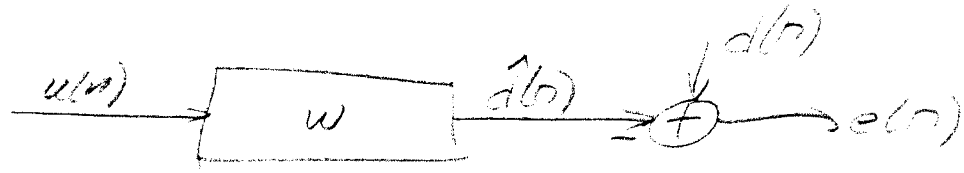
Adaptiv störningsundertryckning, kancellering



Ex	Brusundertryckning i bil, flygplan, mobil, etc	P: Tal + brus1 R: brus2 (korr med brus1)
	50 Hz undertryckning i EKG	P: EKG inkl. 50 Hz R: 50 Hz signal från väggen
	Ekosläckning i tex telefon	P: T2 + T1 retur R: T1
	Line enhancer	P: Färgat brus + periodisk sig R: Som P men delayad Δ
	Notch filter	P: Signal R: $c \cdot \cos \omega t$
	Högpassfilter	P: Signal R: 1

Flera R går också bra.

Filtrering



Wienerfilter (MS)

$$J = E[e^2(n)] = E\Big[\Big(d(n) - \sum_{k=0}^{M-1} w(k)u(n-k)\Big)^2\Big]$$

$$= \underbrace{E[d^2(n)]}_{r_d(0)} - 2\sum_{k=0}^{M-1} w(k)\underbrace{E[d(n)u(n-k)]}_{r_{du}(k)} + \sum_{k_1=0}^{M-1} w(k_1)\sum_{k_2=0}^{M-1} w(k_2)\underbrace{E[u(n-k_1)u(n-k_2)]}_{r_u(k_2-k_1)}$$

$$\frac{\partial J}{\partial w(l)} = -2E[d(n)u(n-l)] + 2\sum_{k=0}^{M-1} w(k)E[u(n-k)u(n-l)] = 0 \qquad l = 0..M-1$$

$$\Rightarrow \; -2r_{du}(l) + 2\sum_{k=0}^{M-1} w(k)r_u(l-k) = 0 \qquad l = 0..M-1$$

$$\sum_{k=0}^{M-1} w(k)r_u(l-k) = r_{du}(l) \qquad l = 0..M-1$$

$$\underline{R_u w = r_{du} \quad (\text{alt } p)}$$

där $r_u(l-k) = E[u(n-k)u(n-l)]$

och $r_{du}(l) = E[d(n)u(n-l)] = E[u(n-l)d(n)] = p(-l)$

LS-filter

$$\mathcal{E} = \sum_{i=i_1}^{i_2} e^2(n) = \sum_{i=i_1}^{i_2}\Big(d(n) - \sum_{k=0}^{M-1} w(k)u(n-k)\Big)^2$$

$$= \underbrace{\sum_{i=i_1}^{i_2} d^2(n)}_{\phi_d(0)} - 2\sum_{k=0}^{M-1} w(k)\underbrace{\sum_{i=i_1}^{i_2} d(n)u(n-k)}_{\phi_{du}(k)} + \sum_{k_1=0}^{M-1} w(k_1)\sum_{k_2=0}^{M-1} w(k_2)\underbrace{\sum_{i=i_1}^{i_2} u(n-k_1)u(n-k_2)}_{\phi_u(k_2,k_1)}$$

$$\frac{\partial \mathcal{E}}{\partial w(l)} = -2\sum_{i=i_1}^{i_2} d(n)u(n-l) + 2\sum_{k=0}^{M-1} w(k)\sum_{i=i_1}^{i_2} u(n-k)u(n-l) = 0 \qquad l = 0..M-1$$

$$\Rightarrow \; \sum_{k=0}^{M-1} w(k)\,\Phi_u(l,k) = \Phi_{du}(l) \qquad l = 0..M-1$$

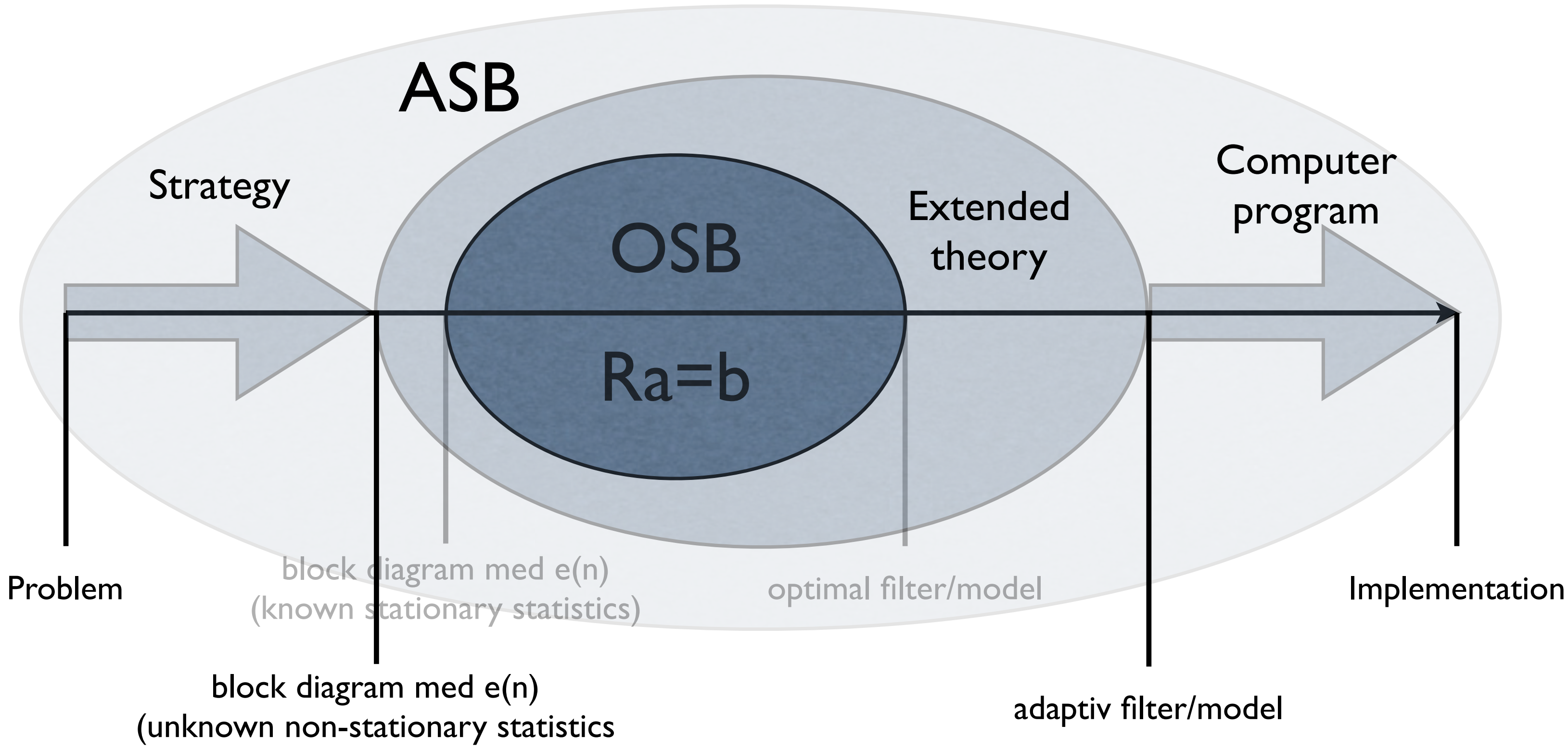
$$\underline{\Phi_u w = \Phi_{du} \quad (\text{alt } z)}$$

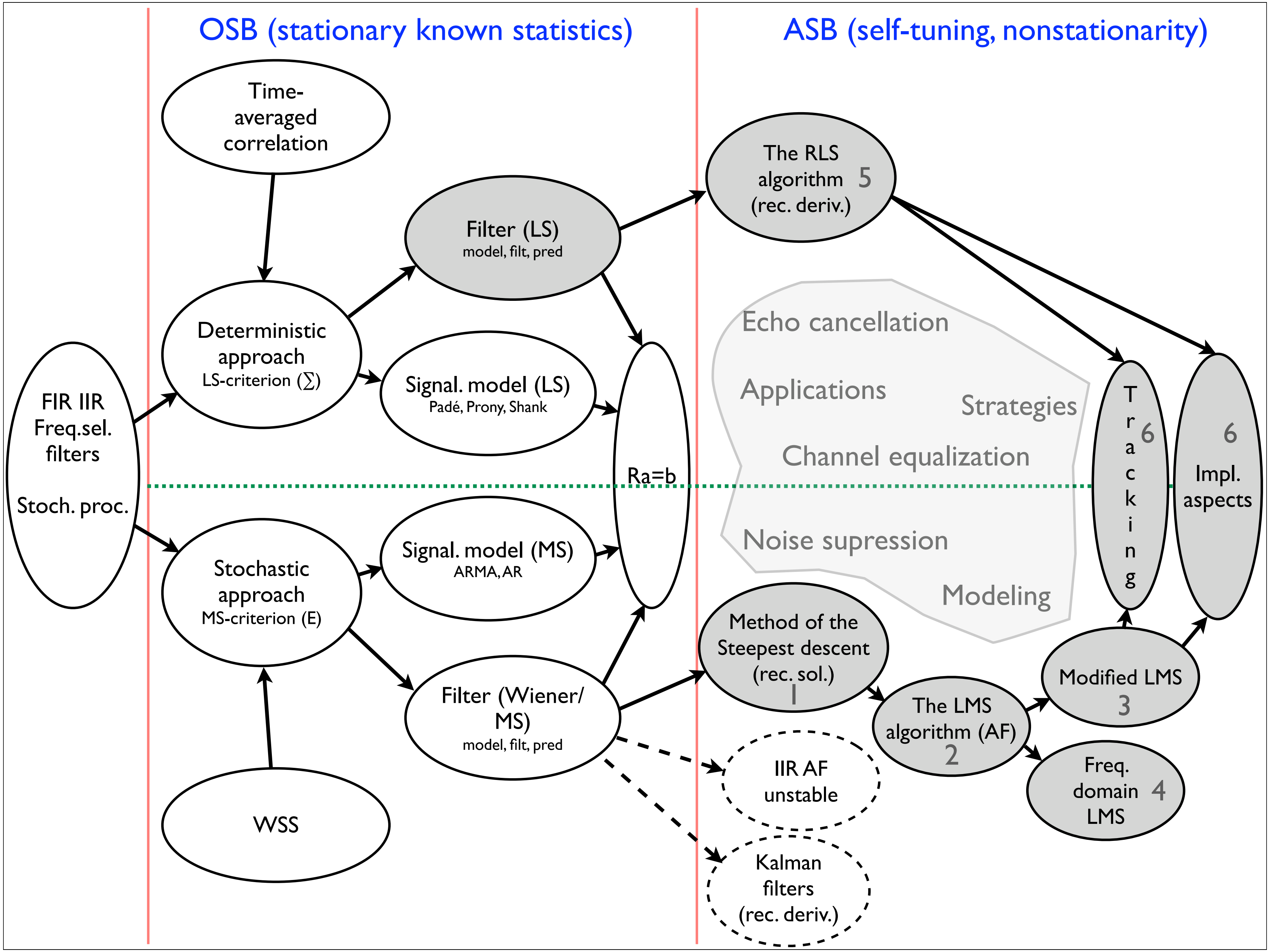
där $\Phi(k,l) = \sum_{i=i_1}^{i_2} u(n-k)u(n-l)$

och

$$\Phi_{du}(l) = \sum_{i=i_1}^{i_2} d(n)u(n-l) = \sum_{i=i_1}^{i_2} u(n-l)d(n) = z(-l)$$

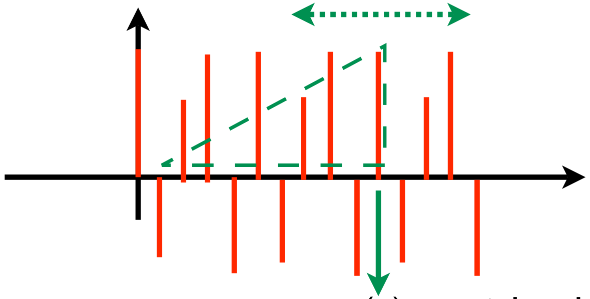
What do you learn in ASB?



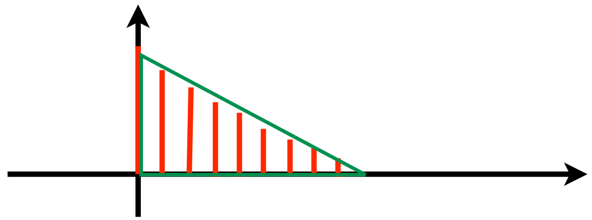

OSB (stationary known statistics)
ASB (self-tuning, nonstationarity)
Time-averaged correlation
The RLS algorithm (rec. deriv.) 5
Filter (LS) model, filt, pred
Deterministic approach LS-criterion (∑)
Signal. model (LS) Padé, Prony, Shank
FIR IIR Freq.sel. filters
Stoch. proc.
Ra=b
Echo cancellation
Applications
Strategies
Channel equalization
Noise supression
Modeling
Tracking 6
6 Impl. aspects
Stochastic approach MS-criterion (E)
Signal. model (MS) ARMA, AR
Filter (Wiener/MS) model, filt, pred
Method of the Steepest descent (rec. sol.) 1
The LMS algorithm (AF) 2
Modified LMS 3
Freq. domain LMS 4
IIR AF unstable
Kalman filters (rec. deriv.)
WSS

Linear convolution

signal (long)



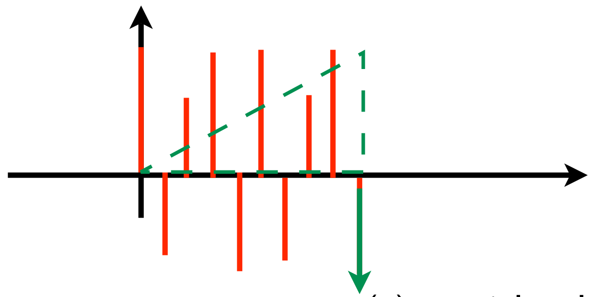
filter (short, length M)



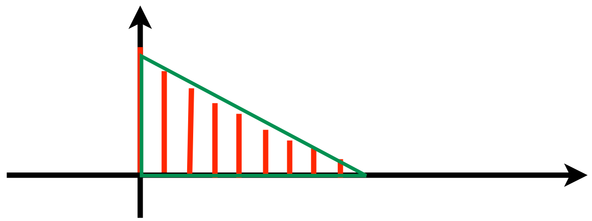
y(n)= weighted sum of u(n-M+1)...u(n)

Circular convolution

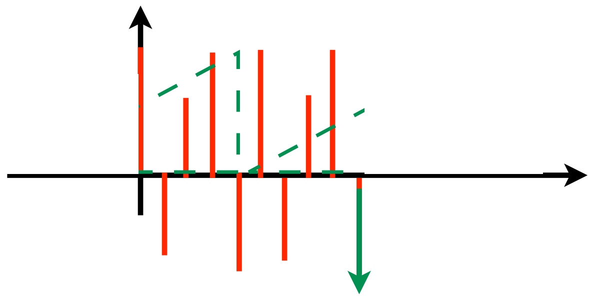
signal 1



signal2/filter



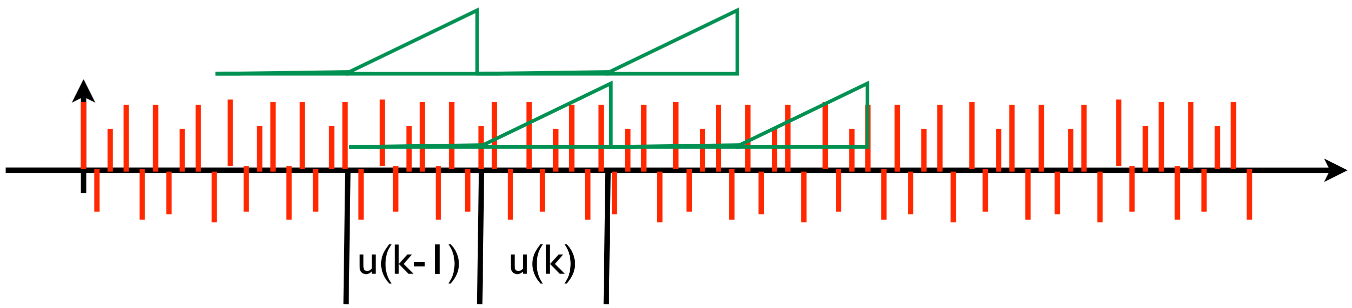
y(n)= weighted sum of u(n-M+1)...u(n)

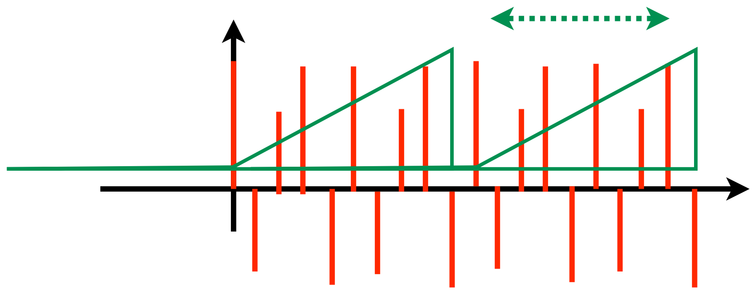


u(n)*w(n)=[. . . . . . . y(n)]

Linear convolution using circular convolution

Use filter of length 2M (block length M)





First M-1 convolutions not OK
Last M+1 convolutions OK

Conclusions

$$\begin{array}{|c|c|}\hline u(k-1) & u(k) \\ \hline \end{array} * \begin{array}{|c|c|}\hline 0 & w(n) \\ \hline \end{array} \longrightarrow \begin{array}{|c|c|}\hline \ldots\ldots & y(k) \\ \hline \end{array}$$

Similary

$$\begin{array}{|c|c|}\hline u(k)^R & u(k-1)^R \\ \hline \end{array} * \begin{array}{|c|c|}\hline 0 & e(k) \\ \hline \end{array} \longrightarrow \begin{array}{|c|c|}\hline \text{grad } J(k) & \ldots\ldots \\ \hline \end{array}$$