Main content ADSP course

- Part A: Stochastic Signal Processing
- **Part B: Adaptive Signal Processing**
- Part C: Array Signal Processing (ASP) (including DOA)
- Part D: Adaptive Array Signal Processing (AASP)

ADSP Part B: Adaptive signal processing

Focus on single channel adaptive algorithms using FIR structure

- Applications Adaptive Algorithms
- Minimum Mean Square Error (MMSE)
- Constrained MMSE
- Least Square (LS)

Fixed weights

- Steepest Gradient Descent (SGD)
- Three LMS variants: NLMS, Complex LMS, Constrained LMS
- Newton
- Recursive Least Squares (RLS) 自回归最小二乘
- Frequency Domain Adaptive Filter (FDAF)
- Summary

上海科技大学
ShanghaiTech University

ADSP Part B: Adaptive signal processing

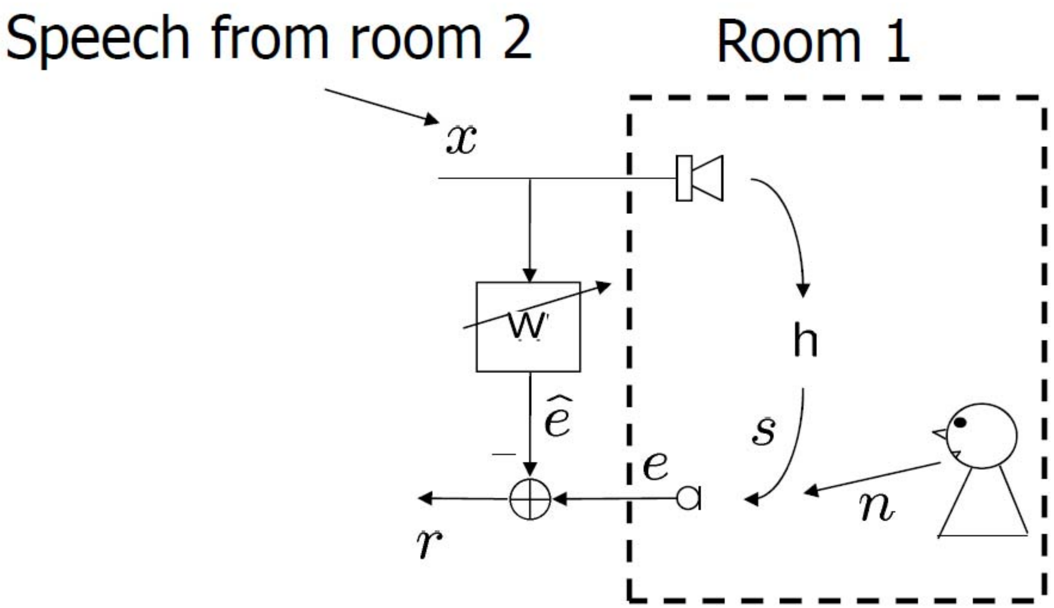
Focus on single channel adaptive algorithms using FIR structure

- Applications Adaptive Algorithms
- Minimum Mean Square Error (MMSE)
- Constrained MMSE
- Least Square (LS)
- Steepest Gradient Descent (SGD)
- Three LMS variants: NLMS, Complex LMS, Constrained LMS
- Newton
- Recursive Least Squares (RLS)
- Frequency Domain Adaptive Filter (FDAF)
- Summary

Three different application classes

1. Signal estimation/ System identification

Example: Acoustic echo cancelling

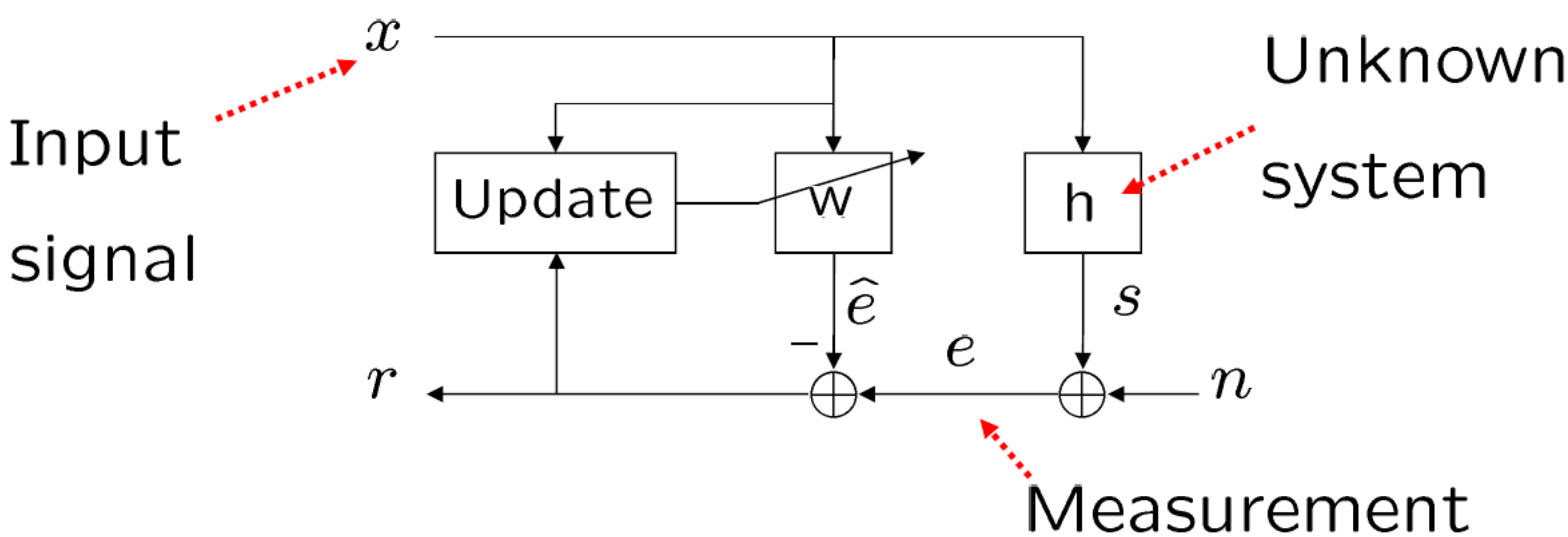


ADSP

Three different application classes

1. Signal estimation/ System identification

Example: Acoustic echo cancelling



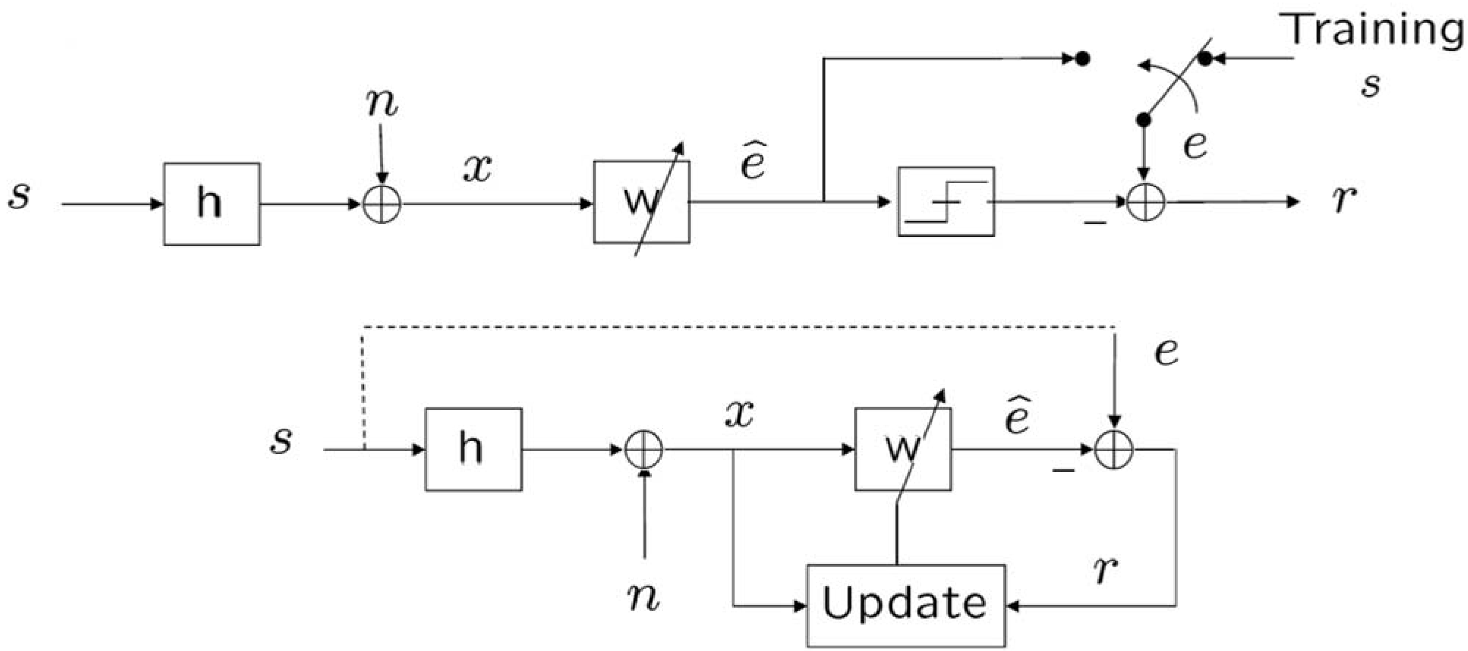
$\hat{e}$ estimate signal $s \Rightarrow$ output: $r\ (\approx n)$

ADSP

Three different application classes

2. Signal correction

Example: equalization



Correct noisy input $x \Rightarrow$ output: $\hat{e}$ $(\approx s)$

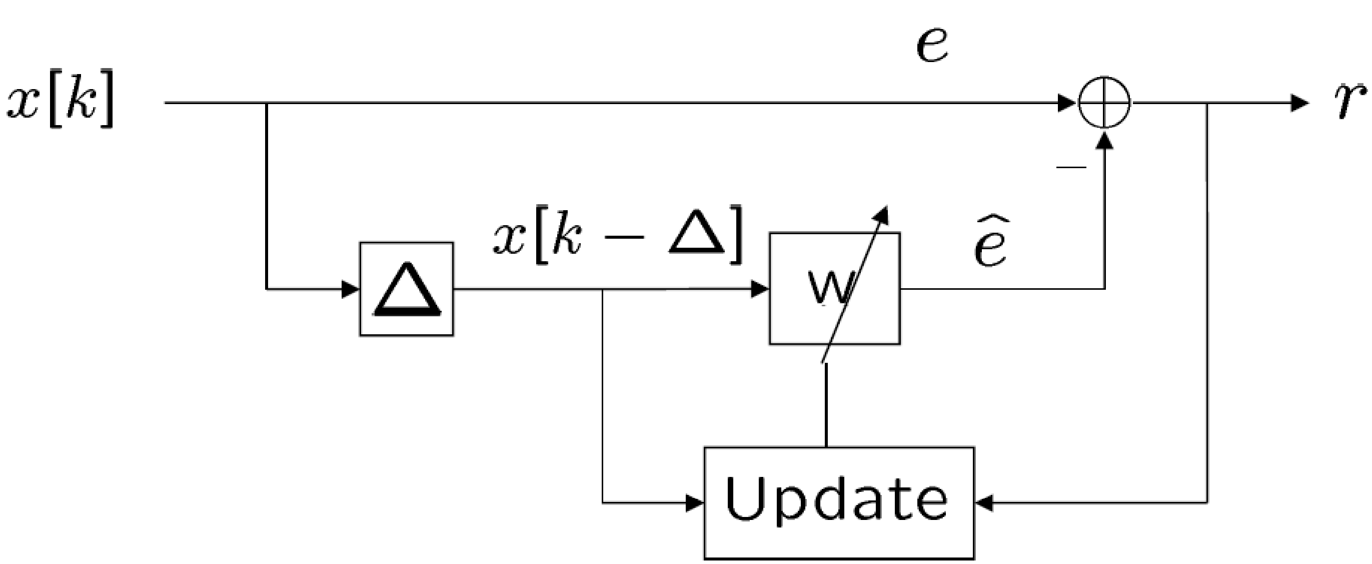
ADSP

Three different application classes

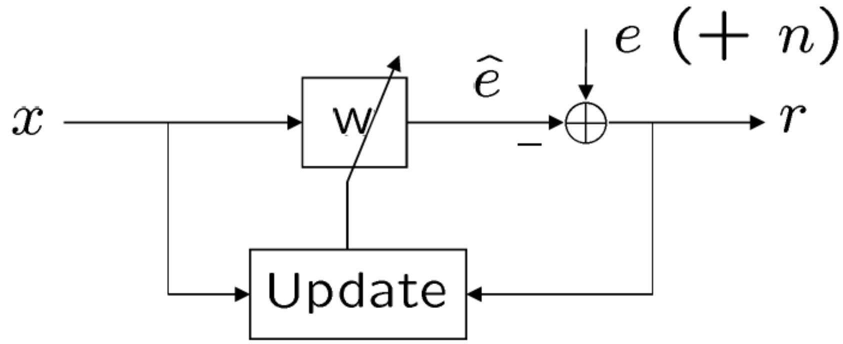
3. Signal prediction

Predict $x[k]$ from $x[k-\Delta]$

$\Rightarrow$ output: $\hat{e}$ $(\approx x[k])$



General model



Notes:

Signals x, e correlated

Noise n not correlated with other signals

Pragmatic choices:
实际的选择

- All signal average zero
- Filter w: FIR

How calculate filter weights w?

- Use quadratic cost function: $J = f(r^2)$
- First fixed weights (MMSE, LS), then adaptive

Part B: Adaptive signal processing

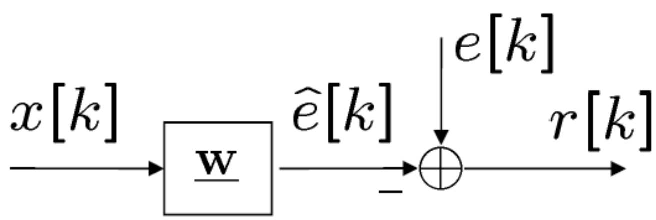
Focus on single channel adaptive algorithms using FIR structure

- Applications Adaptive Algorithms
- Minimum Mean Square Error (MMSE)
- Constrained MMSE
- Least Square (LS)
- Steepest Gradient Descent (SGD)
- Three LMS variants: NLMS, Complex LMS, Constrained LMS
- Newton
- Recursive Least Squares (RLS)
- Frequency Domain Adaptive Filter (FDAF)
- Summary

MMSE

First calculate fixed filter $\underline{\mathbf{w}} = (w_0, w_1, \cdots, w_{N-1})^t$

that Minimizes Mean Square Error (MMSE)



Optimization problem:

Given FIR samples $x[k-i]$ for $i = 0, 1, \cdots, N-1$

$$\underline{\mathbf{w}}_o = \arg\min_{\underline{\mathbf{w}}} \{E\{r^2[k]\}\}$$

MMSE

with $\underline{\mathbf{x}}[k] = (x[k], x[k-1], \cdots, x[k-N+1])^t$

and $\underline{\mathbf{w}} = (w_0, w_1, \cdots, w_{N-1})^t$

MSE:

$$J = E\{r^2[k]\} = E\{(e[k] - \hat{e}[k])^2\}$$

标量

$$= E\{(e[k] - \underline{\mathbf{w}}^t \cdot \underline{\mathbf{x}}[k])(e[k] - \underline{\mathbf{x}}^t[k] \cdot \underline{\mathbf{w}})\}$$

for k

Neglecting time indices: $\underline{w}$ fixed

$$J = E\{e^2\} - \underline{\mathbf{w}}^t E\{\underline{\mathbf{x}} e\} - E\{e \underline{\mathbf{x}}^t\} \underline{\mathbf{w}} + \underline{\mathbf{w}}^t E\{\underline{\mathbf{x}} \underline{\mathbf{x}}^t\} \underline{\mathbf{w}}$$

$$= E\{e^2\} - \underline{\mathbf{w}}^t \underline{\mathbf{r}}_{ex} - \underline{\mathbf{r}}_{ex}^t \underline{\mathbf{w}} + \underline{\mathbf{w}}^t \mathbf{R}_x \underline{\mathbf{w}}$$

MMSE

Autocorrelation $r_x[l] = \mathrm{E}\{x[k]x[k-|l|]\}$

Auto correlation matrix:

$$\mathrm{R}_x = \mathrm{E}\{\underline{\mathbf{x}} \cdot \underline{\mathbf{x}}^t\} = \begin{bmatrix} r_x[0] & r_x[1] & \cdots & r_x[N-1] \\ r_x[1] & r_x[0] & \cdots & r_x[N-2 \\ \vdots & \vdots & \vdots & \vdots \\ r_x[N-1 & r_x[N-2 & \cdots & r_x[0] \end{bmatrix}$$

$$\begin{bmatrix} x[k] \\ x[k-1] \\ \vdots \\ x[k-N+1] \end{bmatrix} \cdot [x[k], x[k-1], \ldots x[k-N+1]]$$

Cross correlation $r_{ex}[l] = \mathrm{E}\{e[k]x[k-l]\}$

Cross correlation vector:

$$\underline{\mathbf{r}}_{ex} = \mathrm{E}\{e \cdot \underline{\mathbf{x}}\} = (r_{ex}[0], r_{ex}[1], \ldots, r_{ex}[N-1])^t$$

$$\mathrm{E}\left\{ e[k] \cdot \begin{bmatrix} x[k] \\ x[k-1] \\ \vdots \\ x[k-N+1] \end{bmatrix} \right\} = \begin{bmatrix} r_{ex}[0] \\ r_{ex}[1] \\ \vdots \\ r_{ex}[N-1] \end{bmatrix}$$

上海科技大学 ShanghaiTech University

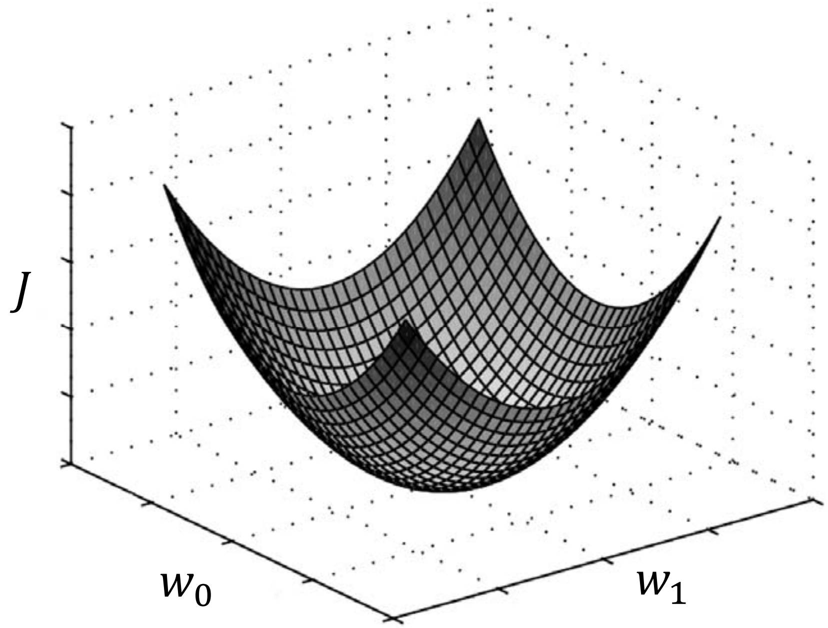
ADSP

MMSE

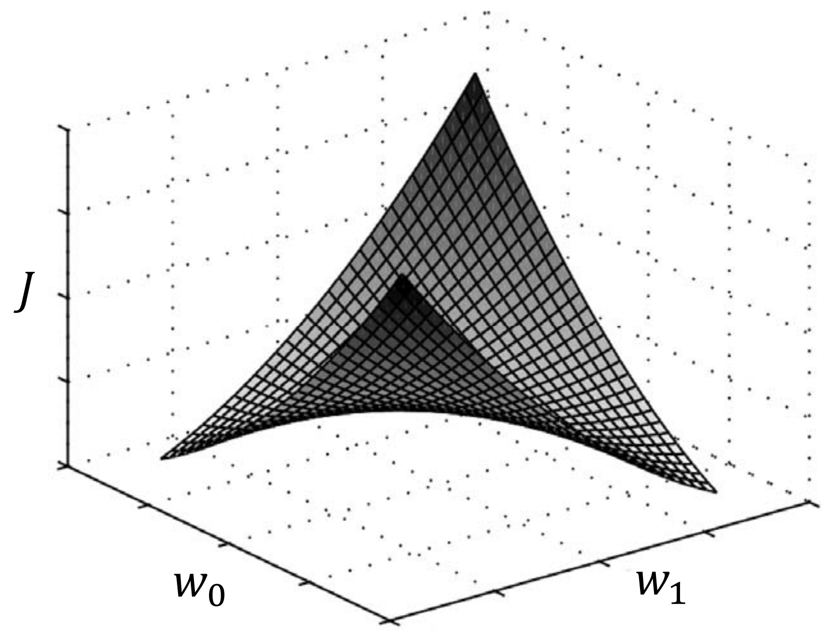
$w^t r_{ex} = r_{ex}^t w$

Cost function $J = E\{e^2\} - \mathbf{w}^t\mathbf{r}_{ex} - \mathbf{r}_{ex}^t\mathbf{w} + \mathbf{w}^t\mathbf{R}_x\mathbf{w}$

In theory if $\boldsymbol{R_x}$ Positive defined, J has unique minimum.



R_x Positive defined



R_x negative defined

ADSP

MMSE

Cost function $J = E\{e^2\} - \underline{\mathbf{w}}^t \underline{\mathbf{r}}_{ex} - \underline{\mathbf{r}}_{ex}^t \underline{\mathbf{w}} + \underline{\mathbf{w}}^t \mathbf{R}_x \underline{\mathbf{w}}$

$\underline{w}_o^t r_{ex} = E\{\hat{d}^2\}$

- $\boldsymbol{R_x}$ always non-negative
- In physical practice, $\boldsymbol{R_x}$ always almost positive

$R_x^T = R_x$, $\underline{a} R_x^{-1} a \geq 0$

After some manipulations we can write:

$$J = \underset{\text{constant}}{E\{e^2\}} - \underbrace{\underline{\mathrm{r}}_{ex}^t \mathrm{R}_x^{-1} \underline{\mathrm{r}}_{ex}}_{\text{和 w 无关} \;\geq 0} + \underbrace{(\mathrm{R}_x \underline{\mathrm{w}} - \underline{\mathrm{r}}_{ex})^t \mathrm{R}_x^{-1} (R_x \underline{\mathrm{w}} - \underline{\mathrm{r}}_{ex})}_{\geq 0}$$

J minimum if $\mathrm{R}_x \underline{\mathrm{w}} = \underline{\mathrm{r}}_{ex}$

$(R_x \underline{w} - \underline{r}_{ex})^T R_x^{-1} (R_x \underline{w} - \underline{r}_{ex})$

$= (\underline{w}^T R_x^T - \underline{r}_{ex}^T) R_x^{-1} (R_x \underline{w} - \underline{r}_{ex})$

$= \underline{w}^T (R_x \underline{w} - \underline{r}_{ex}) - \underline{r}_{ex}^T (\underline{w} - R_x^{-1} \underline{r}_{ex})$

R_x^{-1} also positive

then $\underline{r}_{ex}^{-1} R_x^{-1} \underline{r}_{ex} \geq 0$

$\underline{w}_o = R_x^{-1} r_{ex}$

$\underline{w}_o^t = \underline{r}_{ex}^t R_x^{-1}$

$J - J_{min} = (R_x \underline{w} - \underline{r}_{ex})^t R_x^{-1} R_x R_x^{-1} (R_x \underline{w} - \underline{r}_{ex})$

$= (\underline{w}^t - \underline{r}_{ex}^t R_x^{-1}) R_x (\underline{w} - R_x^{-1} \underline{r}_{ex})$

$= (\underline{w} - \underline{w}_o)^t R_x (\underline{w} - \underline{w}_o)$

$J_{min} = E\{e^2\} - \underline{r}_{ex}^t \underline{w}_o$

$= E\{e^2\} - E\{\hat{e}^2\}$

Part B: Adaptive signal processing

上海科技大学 ShanghaiTech University

ADSP MMSE

Optimum: $$\frac{\mathrm{d}J}{\mathrm{d}\underline{\mathbf{w}}} = \underline{\nabla} = -2\left(\underline{\mathbf{r}}_{ex} - \mathrm{R}_x\underline{\mathbf{w}}\right) = 0$$

$$\Rightarrow \quad \mathrm{R}_x \cdot \underline{\mathbf{w}} = \underline{\mathbf{r}}_{ex}$$ **Normal equations**

$$\boxed{\Rightarrow \quad \underline{\mathbf{w}}_o = \mathrm{R}_x^{-1} \cdot \underline{\mathbf{r}}_{ex} \quad \textbf{Wiener filter}}$$

后续知道有 $r \perp \hat{e}$

Cost function:

$E\{r^2\} = E\{r \cdot (e - \hat{e})\} = E\{r \cdot e\} - E\{r \cdot \hat{e}\}$

<u>*In optimum:*</u> $$J_{min} = J|_{\underline{\mathbf{w}}=\underline{\mathbf{w}}_o} = E\{r^2\} = E\{r \cdot e\}$$

$$= E\{e^2\} - \underline{\mathbf{r}}_{ex}^t \underline{\mathbf{w}}_o$$

$E(\hat{e}^2) = E(\hat{e}(e - r))$

$= E\{\hat{e}e - \hat{e}r\}$

$= E\{\hat{e}e\}$

$= E\{\underline{x}^t\underline{w}_o e\} = \underline{r}_{ex}^t\underline{w}_o$

$$\underline{\mathbf{r}}_{ex}^t \underline{\mathbf{w}}_o = E(\hat{e}^2)$$

$\hat{e} = \underline{w}_o^t \underline{x} = \underline{x}^t \underline{w}_o = e - r$

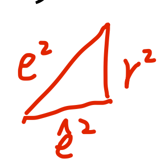
$E(\hat{e}^2) = E\{\underline{x}^t \underline{w}_o (e - r)\}$

$= E\{e\,\underline{x}^t \underline{w}_o\} - E\{\hat{e} \cdot r\}$

$= \underline{r}_{ex}^t \underline{w}_o$

Best case, $\underline{\mathbf{r}}_{ex}^t \neq 0, E\left(\hat{e}^2\right) = E\left(e^2\right), J_{min} = 0$

$J_{min} = E\{e^2\} - E\{\hat{e}^2\}$



$E\{r^2\} = E\{e^2\} - E\{\hat{e}^2\}$

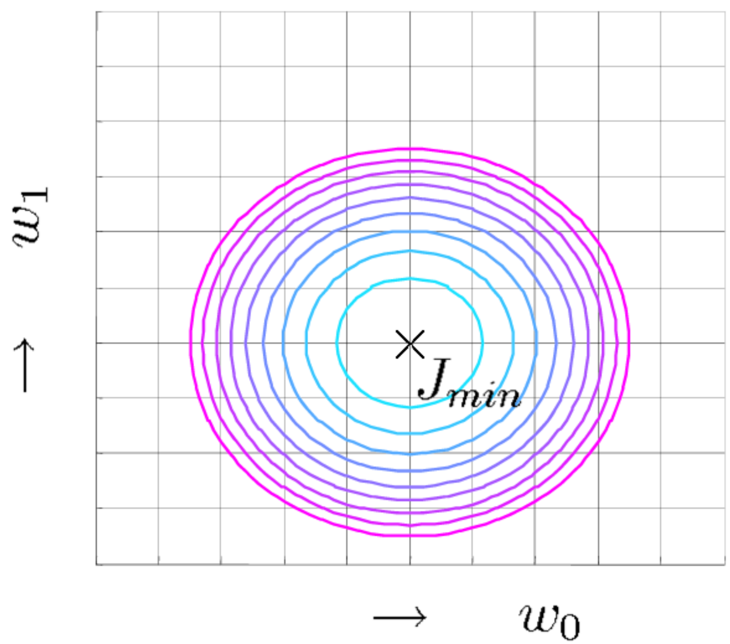
能量对应于强度

MMSE

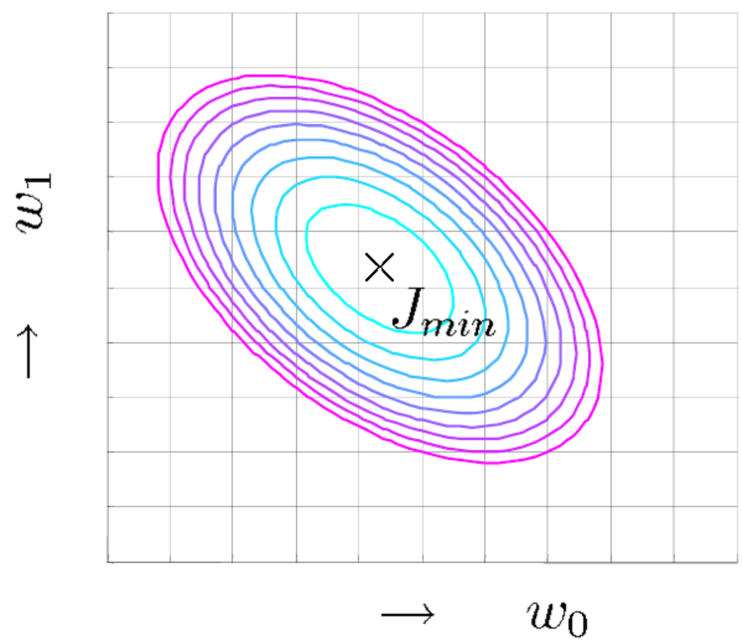
λ_{max} 最大特征值，其所对应的特征向量方向上能量最大

$$\text{Contour plot: } J = J_{min} + (\underline{\mathbf{w}} - \underline{\mathbf{w}}_o)^t \cdot \mathbf{R}_x \cdot (\underline{\mathbf{w}} - \underline{\mathbf{w}}_o)$$

$$\Gamma_x = \frac{\lambda_{max}}{\lambda_{min}} = 1 \text{ (=white noise)}$$



$$\Gamma_x = \frac{\lambda_{max}}{\lambda_{min}} > 1$$



证明技巧: ① $\hat{e} = \underline{w}_o^t \underline{x} = \underline{x}^t \underline{w}_o$ ② $r = e - \hat{e} = e - \underline{w}_o^t \underline{x} = e - \underline{x}^t \underline{w}_o$

Part B: Adaptive signal processing

上海科技大学 ShanghaiTech University

ADSP MMSE

If $w = w_o$, Proof $E\{r\underline{x}\} = \underline{0}$

we have $R_x \underline{w}_o = \underline{r}_{ex}$

$E\{r\underline{x}\} = E\{(e-\hat{e})\underline{x}\} = E\{(e - \underline{x}^t \underline{w}_o)\underline{x}\}$

$= E\{e\underline{x}\} - E\{\underline{x}^t \underline{w}_o \underline{x}\}$ 标量置后

$= E\{e\underline{x}\} - E\{\underline{x}\,\underline{x}^t \underline{w}_o\}$ 是R

$= \underline{r}_{ex} - R_x \underline{w}_o = \underline{0}$

<u>In optimum:</u> $R_x \underline{w} = \underline{r}_{ex}$

Ortogonality principle: $E\{r\underline{x}\} = \underline{0}$

$$\Rightarrow \underline{\nabla}|_{\underline{w}=\underline{w}_o} = E\{r^2\}|_{\underline{w}=\underline{w}_o} = E\{r\cdot(e-\hat{e})\}|_{\underline{w}=\underline{w}_o} = E\{r\cdot e\}$$

$E\{r\hat{e}\} = 0$

from $E\{r\underline{x}\} = \underline{0}$ to proof $E\{r\hat{e}\} = 0$; $E\{r\hat{e}\} = E\{r\cdot \underline{w}_o^t \underline{x}\} = \underline{w}_o^t E\{r\cdot\underline{x}\} = 0$

Geometric interpretation

Optimum:

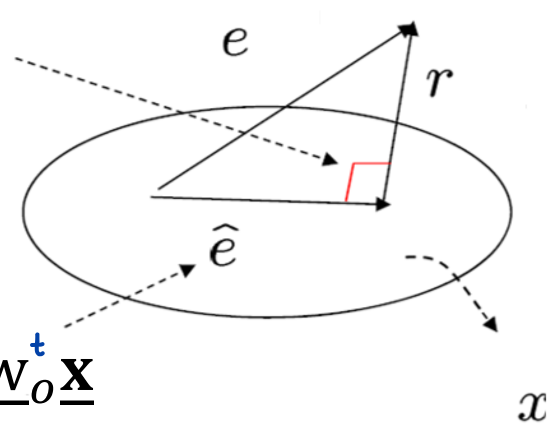
wiener时 $\Rightarrow$ Orthogonal

r的自相关 等于 r与e互相关

ê的自相关 等于 ê与e互相关

$\Rightarrow E\{r^2\} = E\{re\}$

$E\{\hat{e}^2\} = E\{\hat{e}e\}$



$\hat{e} = \underline{w}_o^t \underline{\mathbf{x}}$

$J_{min} = \gamma_r = \gamma_{er}$

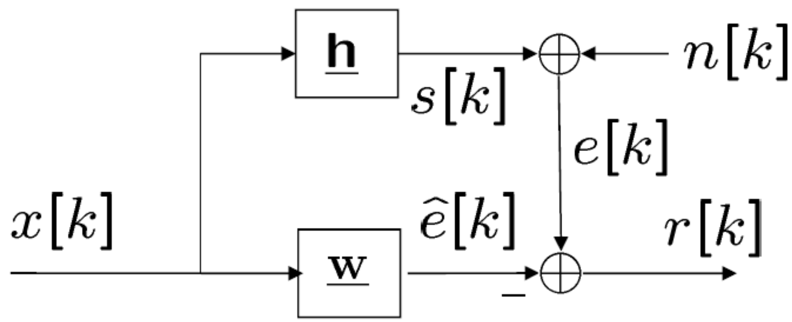
Length:

$J_{min} = E\{r^2\}|_{\underline{w}=\underline{w}_o}$ 损失函数最初的定义

已知输入的延迟信号 x 与参考信号 e, 找出最优 wiener 时 r 与 x, e 的关系。

MMSE

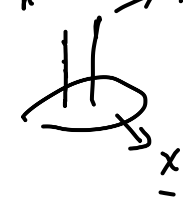
Example: Signal estimation / System identification



理解：

(e-s) 与 (e-ê) 都⊥ x

问题 n? → 仅r吗?



$$\begin{aligned}\underline{\mathbf{r}}_{ex} &= E\{e[k]\underline{\mathbf{x}}[k]\} = E\{\underline{\mathbf{x}}[k]e[k]\} \\ &= E\{\underline{\mathbf{x}}[k]\cdot\underline{\mathbf{x}}^t[k]\underline{\mathbf{h}}\} + E\{\underline{\mathbf{x}}[k]\cdot n[k]\} \\ &= \mathbf{R}_x\cdot\underline{\mathbf{h}} + \underline{\mathbf{0}}\end{aligned}$$

$$\left.\begin{aligned}&\underline{\mathbf{r}}_{ex} = \mathbf{R}_x\cdot\underline{\mathbf{h}} + \underline{\mathbf{0}} \\ &\underline{\mathbf{w}}_o = \mathbf{R}_x^{-1}\cdot\underline{\mathbf{r}}_{ex}\end{aligned}\right\} \Rightarrow \underline{\mathbf{w}}_o = \mathbf{R}_x^{-1}\cdot\mathbf{R}_x\cdot\underline{\mathbf{h}} = \underline{\mathbf{h}}$$

ADSP

Two MMSE variants

- **Complex-MMSE:**

  Setup with complex signals and weights

  Result similar to previous case:

  $$\boxed{\Rightarrow \text{ Optimum: } \underline{\mathbf{w}}_o = \mathbf{R}_x^{-1} \cdot \underline{\mathbf{r}}_{xe^*}}$$

- **Constrained MMSE:**

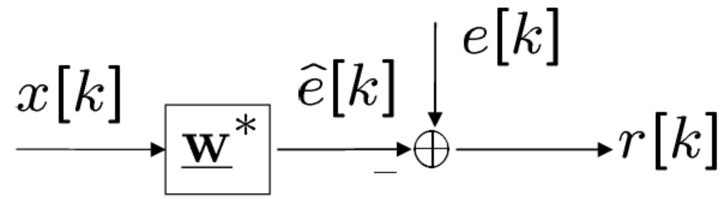
  Setup with set of constraints on weights

ADSP

Complex MMSE

! Complex

$x[k], e[k], w_i$



Input signal vector: $\underline{\mathbf{x}}[k] = (x[k], \cdots, x[k-N+1])^t$

Weight vector: $\underline{\mathbf{w}} = (w_0, \cdots, w_{N-1})^t$

Optimization problem: $\underline{\mathbf{w}}_o = arg \min_{\underline{\mathbf{w}}} \left[E\{|r[k]|^2\}\right]$

MSE: $J = E\{|r[k]|^2\} = E\{r[k] \cdot r^*[k]\}$

$$\begin{aligned} J &= E\{(e[k] - \underline{\mathbf{w}}^h \cdot \underline{\mathbf{x}}[k])(e^*[k] - \underline{\mathbf{x}}^h[k] \cdot \underline{\mathbf{w}})\} \\ &= \sigma_e^2 - \underline{\mathbf{w}}^h \cdot \underline{\mathbf{r}}_{xe^*} - \underline{\mathbf{r}}_{xe^*}^h \cdot \underline{\mathbf{w}} + \underline{\mathbf{w}}^h \cdot \mathbf{R}_x \cdot \underline{\mathbf{w}} \end{aligned}$$

with $\underline{\mathbf{r}}_{xe^*} = E\{\underline{\mathbf{x}}[k]e^*[k]\}$ and $\mathbf{R}_x = \{\underline{\mathbf{x}}[k] \cdot \underline{\mathbf{x}}^h[k]\}$

ADSP

Complex MMSE

After some manipulations we can write:

$$J = \sigma_e^2 - \underline{\mathbf{r}}_{xe^*}^h \mathbf{R}_x^{-1} \underline{\mathbf{r}}_{xe^*} +$$
$$+ (\mathbf{R}_x \cdot \underline{\mathbf{w}} - \underline{\mathbf{r}}_{xe^*})^h \cdot \mathbf{R}_x^{-1} (\mathbf{R}_x \cdot \underline{\mathbf{w}} - \underline{\mathbf{r}}_{xe^*})$$

$\Rightarrow$ Only second part depending on $\underline{\mathbf{w}}$

$\boxed{\Rightarrow \text{ Optimum: } \underline{\mathbf{w}}_o = \mathbf{R}_x^{-1} \cdot \underline{\mathbf{r}}_{xe^*}}$

Other approch: Via (complex) gradient

Assume we can write compex weight as: $\underline{\mathbf{w}} = \underline{\mathbf{a}} + \mathrm{j}\underline{\mathbf{b}}$

$$\underline{\nabla} w(J) = \frac{dJ}{d\underline{\mathbf{a}}} + \mathrm{j}\frac{dJ}{d\underline{\mathbf{b}}} \Leftrightarrow \underline{\nabla} w(J) = 2\frac{dJ}{d\underline{\mathbf{w}}^*} = 2\left(\mathbf{R}_x \cdot \underline{\mathbf{w}} - \underline{\mathbf{r}}_{xe^*}\right)$$

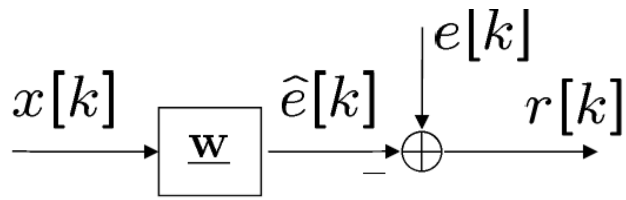
ADSP Part B: Adaptive signal processing

Focus on single channel adaptive algorithms using FIR structure

- Applications Adaptive Algorithms
- Minimum Mean Square Error (MMSE)
- Constrained MMSE
- Least Square (LS)
- Steepest Gradient Descent (SGD)
- Three LMS variants: NLMS, Complex LMS, Constrained LMS
- Newton
- Recursive Least Squares (RLS)
- Frequency Domain Adaptive Filter (FDAF)
- Summary

Constrained MMSE

.



Minimize $E\{r^2[k]\}$ subject to M constraints:

Example: $$\sum_{i=0}^{N-1} w_i = 1 \;\Rightarrow\; (1,\cdots,1)\cdot\begin{pmatrix} w_0 \\ \cdot \\ \cdot \\ w_{N-1} \end{pmatrix} = 1$$

General:

$$\begin{matrix} \underline{\mathbf{c}}_1^t \cdot \underline{\mathbf{w}} = f_1 \\ \vdots \\ \underline{\mathbf{c}}_M^t \cdot \underline{\mathbf{w}} = f_M \end{matrix} \;\Leftrightarrow\; \begin{pmatrix} \underline{\mathbf{c}}_1^t \\ \vdots \\ \underline{\mathbf{c}}_M^t \end{pmatrix} \cdot \underline{\mathbf{w}} = \underline{\mathbf{f}} \;\Leftrightarrow\; \mathbf{C}^t \cdot \underline{\mathbf{w}} = \underline{\mathbf{f}}$$

$M \cdot N$

Constrained MMSE

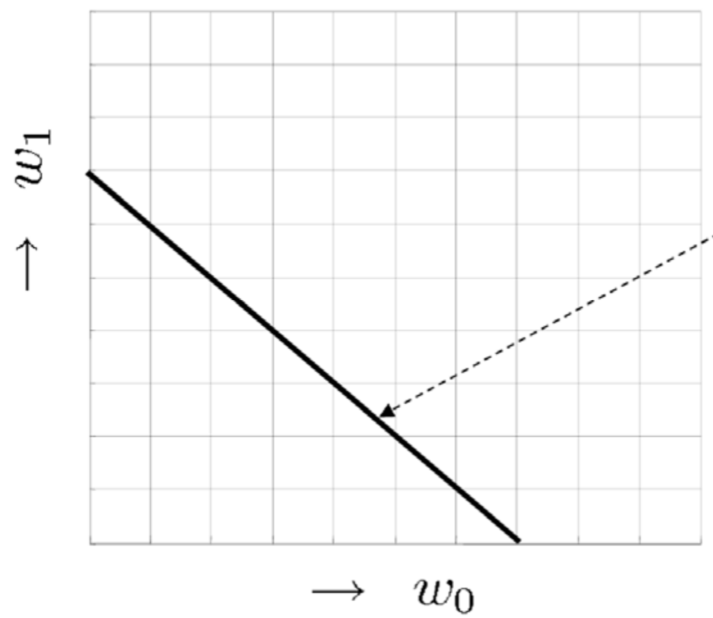
Some notes on solving: $\mathbf{C}^t \cdot \underline{\mathbf{w}} = \underline{\mathbf{f}}$

$M \times 1$ constraint vector: $\underline{\mathbf{f}} = (f_1, \cdots, f_M)^t$

$N \times M$ constraint matrix: $\mathbf{C} = (\underline{\mathbf{c}}_1, \cdots, \underline{\mathbf{c}}_M)$

Note: M independent constraints $\Rightarrow$ $\mathbf{C}$ full rank

$M \leq N$



$\{\underline{\mathbf{w}} : \mathbf{C}^t \cdot \underline{\mathbf{w}} = \underline{\mathbf{f}}\}$

Example of constraint:

$w_0 + w_1 = c$

Constrained MMSE

$$\text{Solution } \mathbf{C}^t \cdot \underline{\mathbf{w}} = \underline{\mathbf{f}} :$$

Case $N = M$:

$$\Rightarrow \quad \underline{\mathbf{w}}^c = \left(\mathbf{C}^t\right)^{-1} \cdot \underline{\mathbf{f}}$$

$$\Rightarrow \quad \textit{No degrees of freedom left for MMSE}$$

Case $N > M$:

$$\Rightarrow \quad \text{Possible solution: } \underline{\mathbf{w}}^c = \left(\mathbf{C}^t\right)^{\dagger} \cdot \underline{\mathbf{f}}$$

$$\textit{Appendix}\text{: Generalized inverse } \left(\mathbf{C}^t\right)^{\dagger} = \mathbf{C} \cdot \left(\mathbf{C}^t \cdot \mathbf{C}\right)^{-1}$$

$$\Rightarrow \quad N - M \ \textit{degrees of freedom left over for MMSE}$$

$N < M \quad \Rightarrow$ Conflicting solution

(choose e.g. minimum norm soluton)