3 Optimal and Adaptive Filtering

3.4: Examples

Optimal and Adaptive Filtering

1. **Wiener-Hopf filter**
   - Minimum Mean Square Error Estimation
   - The Wiener-Hopf solution
2. **Linear prediction**
   - The Wiener-Hopf filter as a predictor
   - Linear prediction for signal coding
3. **Adaptive filtering**
   - Steepest descent
   - Least Mean Square approach
4. **Applications of optimal and adaptive filtering**
   - …

Examples of Optimal & Adaptive Filtering

1. **Affine predictor**
   - Comparison between linear and affine prediction for non-zero mean signals
2. **Wiener-Hopf solution for highly correlated data**
   - Avoiding the use of close to singular matrices
3. **Short term / Long term correlation**
   - Embedding a signal into noise

Affine predictor (I)

3.3

- When we want to predict a non-zero mean correlated signal ($E\big[x[n]\big] = m$) the affine predictor is a better alternative than the linear predictor, since it results in a smaller prediction error power. For the case of order 1, the equations are given by:

- **Linear predictor**: $\hat{x}_p[n] = h_p x[n-1]$
- **Affine predictor:** $\hat{x}_a[n] = h_a x[n-1] + b = [h_a \quad b] \begin{bmatrix} x[n-1] \\ 1 \end{bmatrix}$

a) Obtain the equations of the first order linear predictor

Example of non-zero mean signal

a) $\hat{x}[n] = h_p x[n-1]$ $\Rightarrow$ 1-D FILTER : $\underline{h} = h_p$

$\Rightarrow$ OBSERV. : $x[n-1]$

$\rightarrow$ REFENCE : $x[n]$

$\underline{\underline{R}}_x \cdot \underline{h}_{OPT} = \underline{r}_{xd}$.

$\underline{\underline{R}}_x = r_x(0)$

$\underline{h}_{OPT} = h_{p\,OPT}$

Affine predictor (II)

a) Obtain the equations of the first order linear predictor

$$\underline{r}_{xd} = E\{\underline{x}[n]\, d[n]\} = E\{x[n-1]\, x[n]\} = r_x[-1] = r_x[1]$$

$$\underline{\underline{R}}_x \cdot \underline{h}_{opt} = \underline{r}_{xd} \Rightarrow r_x[0] \cdot h_{0\,opt} = r_x[1] \Rightarrow h_{0\,opt} = \frac{r_x[1]}{r_x[0]}$$

$$\varepsilon = r_d(0) - \underline{h}_{opt}^T \cdot \underline{r}_{xd} = r_d(0) - \underline{h}_{opt}^T \cdot \underline{\underline{R}}_x \cdot \underline{h}_{opt}$$

$$\varepsilon = r_x(0) - \frac{r_x[1]}{r_x[0]} \cdot r_x[0] \cdot \frac{r_x[1]}{r_x[0]} = r_x[0] - \frac{r_x^2[1]}{r_x[0]} \Rightarrow$$

$$\varepsilon = r_x[0]\left[1 - \frac{r_x^2[1]}{r_x^2[0]}\right] \Rightarrow \varepsilon = r_x[0]\left[1 - \rho^2\right]$$

Affine predictor (III)

3.3

a) Obtain the equations of the first order linear predictor

WHICH IS THE MEAN OF THE ERROR?

$$E\{e[n]\} = E\{d[n] - \hat{x}[n]\} = E\{x[n] - \rho x[n-1]\} =$$

$$= E\{x[n]\} - \rho E\{x[n-1]\} = (1-\rho)\cdot m_x \longrightarrow \text{ONLY } E\{e[n]\} = \emptyset \text{ IF } \rho = 1.$$

b) Obtain the equations of h_a and b, from the minimization of $E[e[n]^2] = E[(x[n] - \hat{x}_a[n])^2]$.

b) $\hat{x}_a[n] = h_a x[n-1] + b = \begin{bmatrix} h_a & b \end{bmatrix} \begin{bmatrix} x[n-1] \\ 1 \end{bmatrix}$

THEN: $\underline{h}^T = \begin{bmatrix} h_a & b \end{bmatrix}$ $\quad \underline{x}^T[n] = \begin{bmatrix} x[n-1] & 1 \end{bmatrix}$ $\quad d[n] = x[n]$

Affine predictor (IV)

3.3

b) Obtain the equations of h_a and b, from the minimization of $E[e[n]^2] = E[(x[n] - \hat{x}_a[n])^2]$.

$$\nabla_{\underline{h}} E\{e^2[n]\} = [e[n] = [n] - h_a x[n-1] - b] =$$

$$= \nabla_{\underline{h}} E\{[x[n] - h_a x[n-1] - b][x[n] - h_a x[n-1] - b]\} =$$

$$= \nabla_{\underline{h}} E\{x^2[n] - h_a x[n]x[n-1] - x[n]\cdot b - h_a x[n-1]x[n] +$$

$$+ h_a^2 x^2[n-1] + h_a x[n-1]b - b x[n] + b h_a x[n-1] + b^2\} = \underline{0}$$

$$\frac{\partial}{\partial h_a} E\{e^2[n]\} = E\{-x[n-1]x[n] - x[n-1]x[n] + 2h_a x^2[n-1]^2 +$$

$$+ x[n-1]\cdot b + b\,x[n-1]\} = -r_x[1] - r_x[1] + 2h_a r_x[0] +$$

$$+ \mu_x b + b\,\mu_x = 2h_a r_x[0] - 2r_x[1] + 2\mu_x b = 0 \quad (1)$$

Affine predictor (V)

b) Obtain the equations of h_a and b, from the minimization of $E[e[n]^2] = E[(x[n] - \hat{x}_a[n])^2]$.

$$\frac{\partial}{\partial b} E\{e^2[n]\} = E\{-x[n] + h_a x[n-1] - x[n] + h_a x[n-1] + 2b\} = 0$$

$$-m_x + h_a m_x - m_x + h_a m_x + 2b = 2h_a m_x - 2m_x + 2b = 0 \quad (2)$$

$$(1) \quad h_a r_x[0] - r_x[1] + m_x b = 0$$

$$(2) \quad h_a m_x - m_x + b = 0$$

$$\boxed{h_{a_{OPT}} = \frac{r_x[1] - m_x^2}{r_x[0] - m_x^2}}$$

$$\boxed{b_{OPT} = \frac{r_x[0] - r_x[1]}{r_x[0] - m_x^2} m_x}$$

Affine predictor (VI)

3.3

b) Obtain the equations of h_a and b, from the minimization of $E[e[n]^2] = E[(x[n] - \hat{x}_a[n])^2]$.

WHICH IS THE MEAN OF THE ERROR?

$$E\{e[n]\} = E\{d[n] - \hat{x}[n]\} = E\{x[n] - h_{a\,OPT}\, x[n-1] - b_{OPT}\} =$$

$$= E\left\{x[n] - \frac{r_x[1] - m_x^2}{r_x[0] - m_x^2}\, x[n-1] - \frac{r_x[0] - r_x[1]}{r_x[0] - m_x^2}\, m_x\right\} =$$

$$= m_x - \frac{r_x[1] - m_x^2}{r_x[0] - m_x^2}\, m_x - \frac{r_x[0] - r_x[1]}{r_x[0] - m_x^2}\, m_x =$$

$$= m_x - \frac{r_x[1] - m_x^2 + r_x[0] - r_x[1]}{r_x[0] - m_x^2}\, m_x = m_x - m_x = \emptyset$$

Affine predictor (VII)

3.3

b) Obtain the equations of h_a and b, ~~from the minimization of~~ $\cancel{E[e[n]^2] = E[(x[n] - \hat{x}_a[n])^2]}$.

FROM THE WIENER-HOPF FINAL RESULT

$$\underline{\underline{R}}_x \underline{h}_{OPT} = \underline{r}_{xd} \Rightarrow \underline{\underline{R}}_x = E\left\{ \underline{x}[n]\, \underline{x}^T[n] \right\} = E\left\{ \begin{bmatrix} x[n-1] \\ 1 \end{bmatrix} \begin{bmatrix} x[n-1] & 1 \end{bmatrix} \right\} =$$

$$= E\left\{ \begin{bmatrix} x^2[n-1] & x[n-1] \\ x[n-1] & 1 \end{bmatrix} \right\} = \begin{bmatrix} r_x[0] & m_x \\ m_x & 1 \end{bmatrix}$$

$$\underline{r}_{xd} = E\left\{ \underline{x}[n]\, d[n] \right\} = E\left\{ \begin{bmatrix} x[n-1] \\ 1 \end{bmatrix} x[n] \right\} = \begin{bmatrix} r_x[1] \\ m_x \end{bmatrix}$$

$$\begin{bmatrix} r_x[0] & m_x \\ m_x & 1 \end{bmatrix} \begin{bmatrix} h_a \\ b \end{bmatrix} = \begin{bmatrix} r_x[1] \\ m_x \end{bmatrix} \Rightarrow \begin{array}{ll} (1) & h_a r_x[0] - r_x[1] + m_x b = 0 \\ (2) & h_a m_x - m_x + b = 0 \end{array}$$

Affine predictor (VIII)

3.3

c) For the previous case, what is the expression of the minimum prediction error power?.

$$\mathcal{E} = r_d[0] - \underline{h}_{opt}^T \cdot \underline{r}_{xd} \Rightarrow \mathcal{E} = r_x[0] - \begin{bmatrix} ha_{opt} & b_{opt} \end{bmatrix} \begin{bmatrix} r_x[1] \\ \mu_x \end{bmatrix}$$

$$ha_{opt} = \frac{r_x[1] - \mu_x^2}{r_x[0] - \mu_x^2} = \frac{c_x[1]}{c_x[0]}$$

$$b_{opt} = \frac{r_x[0] - r_x[1]}{r_x[0] - \mu_x^2}\mu_x = \frac{c_x[0] - c_x[1]}{c_x[0]}\mu_x$$

$$\mathcal{E} = c_x[0] + \mu_x^2 - \frac{c_x[1]}{c_x[0]}\left(c_x[1] + \mu_x^2\right) - \frac{c_x[0] - c_x[1]}{c_x[0]}\mu_x^2 =$$

$$\mathcal{E} = c_x[0] + \mu_x^2 - \frac{c_x^2[1]}{c_x[0]} - \frac{c_x[1]}{c_x[0]}\mu_x^2 - \mu_x^2 + \frac{c_x[1]}{c_x[0]}\mu_x^2 =$$

$$\mathcal{E} = c_x[0] - \frac{c_x^2[1]}{c_x[0]} \Rightarrow \mathcal{E} = c_x[0]\left[1 - \frac{c_x^2[1]}{c_x^2[0]}\right]$$

Affine predictor (IX)

d) Suppose that we want to implement the first order affine predictor adaptively. Find the equations of the LMS from the instantaneous estimation (stochastic approximation) of the gradient?.

$$\underline{h}[n+1] = \underline{h}[n] - \frac{1}{2}\mu \nabla E\{e^2[n]\}\Big|_{\underline{h}[n]}$$

$$\underline{h}[n+1] = \underline{h}[n] + \mu\, \underline{x}[n]\, e[n] \quad \Rightarrow \quad \underline{x}[n] = \begin{bmatrix} x[n-1] \\ 1 \end{bmatrix}$$

$$\underline{h}[n+1] = \underline{h}[n] + \mu \begin{bmatrix} x[n-1] \\ 1 \end{bmatrix} e[n] \Rightarrow \begin{cases} h_a[n+1] = h_a[n] + \mu\, x[n-1]\, e[n] \\ b[n+1] = b[n] + \mu\, e[n] \end{cases}$$

Affine predictor (X)

3.3

e) Which interval of μ values assures the convergence?

e) $0 < \mu < \frac{2}{\lambda_{MAX}}$ $\Rightarrow$ $\underline{\underline{R}}_x = E\{\underline{x}[n] \cdot \underline{x}^T[n]\}$ $\Rightarrow$

$\underline{\underline{R}}_x = \begin{bmatrix} r_x[0] & \mu_x \\ \mu_x & 1 \end{bmatrix}$ $\Rightarrow$ A MORE RESTRICTIVE BOUND IS USUALLY ADOPTED

$\lambda_{MAX} \leq \sum \lambda_i = \text{Trace}[\underline{\underline{R}}_x] \Rightarrow \lambda_{MAX} \leq r_x[0] + 1$

$0 < \mu < \frac{2}{r_x[0]+1}$ OR EVEN $0 < \mu < \frac{2\alpha}{r_x[0]+1}$ WITH $0 < \alpha < 1$

Examples of Optimal & Adaptive Filtering

3.3

1. **Affine predictor**
   - Comparison between linear and affine prediction for non-zero mean signals
2. **Wiener-Hopf solution for highly correlated data**
   - Avoiding the use of close to singular matrices
3. **Short term / Long term correlation**
   - Embedding a signal into noise

W-H solution for high correlated data (I)

3.3

- Wiener filter architectures have some drawbacks when the data x[n] is highly correlated. In such case, the autocorrelation matrix $\mathbf{R} = E[\mathbf{x}[n]\mathbf{x}[n]^H]$ be singular and not have inverse or may yield to very high values in the optimal vector coefficients h[n] (high values could degenerate the performance of the algorithm, especially if we use finite arithmetic processors). In order to solve the problem of designing the filter coefficients h[n] from the mentioned data and from a training sequence or reference d[n], we need to introduce a penalization factor in the adaptation equation of the vector h[n]. This can be expressed as the minimization of the following objective function:

$$\mathbf{h} = \arg\min_{\mathbf{h}} \{\xi(\mathbf{h}) = E[|e[n]|^2] + \alpha\|\mathbf{h}\|^2\}$$

where α is a real scalar and constant along time and the error is $e[n] = d[n] - \mathbf{h}^T\mathbf{x}[n]$.

a) Find the expression of the optimum filter that minimizes the objective function ξ(h) in stationary conditions

a) $\nabla \xi(\underline{h}) = \nabla\left(E\{e^2[n]\} + \alpha|\underline{h}|^2\right) = \underline{\emptyset}$ $\qquad e[n] = d[n] - \underline{h}^T\underline{x}[n]$

$\nabla_{\underline{h}} E\{(d[n] - \underline{h}^T\underline{x}[n])^T(d[n] - \underline{h}^T\underline{x}[n]) + \alpha\, \underline{h}^T\cdot\underline{h}\} = \underline{\emptyset}$

$E\{-2\,\underline{x}[n]\,d[n] + 2\,\underline{x}[n]\,\underline{x}^T[n]\cdot\underline{h} + \alpha\, 2\,\underline{h}\} = \underline{\emptyset}$

W-H solution for high correlated data (II)

3.3

b) Discuss the role of constant α in the solution found in previous section. Under what conditions would such constant adopt a positive, negative, or zero value?

b) α CAN BE RELATED TO THE POWER OF A WHITE NOISE THAT IS ADDED TO THE SIGNAL $x[n]$

→ $\alpha > 0$ ⇒ ENSURES THE EXISTENCE OF $\underline{\underline{R}}_x^{-1}$

⇒ PENALYZES $|\underline{h}|^2 >>$ → GOOD FOR IMPLEMENTATIONS WITH FINITE ARITHMETIC

→ $\alpha = \emptyset$ ⇒ CONVENCIONAL SOLUTION

→ $\alpha < \emptyset$ ⇒ NO SENSE : $\lambda \notin \mathbb{R}^+$

W-H solution for high correlated data (III)

3.3

c) Obtain the equation of the adaptive filter based on the LMS (i.e., on the stochastic or instantaneous gradient of the objective function $\xi(\mathbf{h})$).

c) $\underline{h}^{n+1} = \underline{h}^{n} - \frac{1}{2}\mu \nabla_{\underline{h}} (\hat{\xi}(\underline{h}))\Big|_{\underline{h}^n} = \left[\underline{\underline{R}}_x \simeq \underline{x}[n]\,\underline{x}^T[n]\ ;\ \underline{r}_{xd} \simeq \underline{x}[n]\,d[n]\right] =$

$\nabla_{\underline{h}} \xi(\underline{h}) = -2\underline{r}_{xd} + 2\underline{\underline{R}}_x \cdot \underline{h} + 2\alpha\underline{h} \Rightarrow$

$\nabla_{\underline{h}} \hat{\xi}(\underline{h}) = -2\,\underline{x}[n]\,d[n] + 2\,\underline{x}[n]\,\underline{x}^T[n]\,\underline{h} + 2\alpha\,\underline{h} =$

$\underline{h}^{n+1} = \underline{h}^{n} + \mu\,\underline{x}[n]\left(d[n] - \underline{x}^T[n]\cdot\underline{h}^n\right) - \mu\alpha\,\underline{h}^n$

$\underline{h}^{n+1} = (1-\mu\alpha)\,\underline{h}^n + \mu\,\underline{x}[n]\,e[n]$

W-H solution for high correlated data (IV)

3.3

d) Prove that the adaptive filter found in the previous section converges in mean to the solution obtained in the first section a).

d) $E\{\underline{h}^{n+1}\} = E\{\underline{h}^{n} + \mu\, \underline{x}[n]\left(d[n] - \underline{x}^{T}[n]\cdot \underline{h}^{n}\right) - \mu\alpha\, \underline{h}^{n}\} =$

$$E\{\underline{h}^{n+1}\} = E\{\underline{h}^{n}\} + \mu\, E\{\underline{x}[n]\,d[n]\} - \mu\, E\{\underline{x}[n]\,\underline{x}^{T}[n]\,\underline{h}^{n}\} - E\{\mu\alpha\,\underline{h}^{n}\}$$

$$E\{\underline{h}^{n+1}\} = E\{\underline{h}^{n}\} + \mu\, \underline{r}_{xd} - \mu\, \underline{\underline{R}}_{x}\cdot E\{\underline{h}^{n}\} - \mu\alpha\, E\{\underline{h}^{n}\} =$$

$$E\{\underline{h}^{n+1}\} = E\{\underline{h}^{n}\} + \mu\left(\underline{\underline{R}}_{x} + \alpha\,\underline{\underline{I}}\right)\underline{h}_{opt} - \mu\left(\underline{\underline{R}}_{x} + \alpha\,\underline{\underline{I}}\right)E\{\underline{h}^{n}\}$$

$$E\{\underline{h}^{n+1}\} - \underline{h}_{opt} = E\{\underline{h}^{n}\} - \underline{h}_{opt} - \mu\left(\underline{\underline{R}}_{x} + \alpha\,\underline{\underline{I}}\right)\left(E\{\underline{h}^{n}\} - \underline{h}_{opt}\right)$$

$$E\{\underline{h}^{n+1} - \underline{h}_{opt}\} = \left[\underline{\underline{I}} - \mu\left(\underline{\underline{R}}_{x} + \alpha\,\underline{\underline{I}}\right)\right]\left(E\{\underline{h}^{n}\} - \underline{h}_{opt}\right)$$

W-H solution for high correlated data (V)

e) Find the limits of the step-size μ in which the adaptive filter converges to the desired solution.

f) Determine an upper-bound of the convergence time for the previous adaptive algorithm

e) $0 < \mu < \frac{2}{\lambda_{MAX}} \Rightarrow 0 < \mu < \frac{2}{\lambda_{MAX}(\underline{\underline{R}}_x + \alpha \underline{\underline{I}})}$

$$\left.\begin{aligned} \underline{\underline{R}}_x \cdot \underline{v} &= \lambda \underline{v} \\ \alpha \underline{\underline{I}} \cdot \underline{v} &= \alpha \underline{v} \end{aligned}\right\} \quad (\underline{\underline{R}}_x + \alpha \underline{\underline{I}}) \underline{v} = (\lambda + \alpha) \underline{v}$$

$$\lambda_{MAX}(\underline{\underline{R}}_x + \alpha \underline{\underline{I}}) = \lambda_{MAX}(\underline{\underline{R}}_x) + \alpha \qquad 0 < \mu < \frac{2}{\lambda_{MAX} + \alpha}$$

f) $N_{ITER} \propto -\ln \varepsilon \frac{\lambda_{MAX}}{\lambda_{MIN}} \Rightarrow N_{ITER} \propto -\ln \varepsilon \frac{\lambda_{MAX} + \alpha}{\lambda_{MIN} + \alpha}$

W-H solution for high correlated data (V)

g) Now suppose that $\mathbf{h}[n]$ has only two coefficients $\mathbf{h}[n] = [h_0[n] \;\; h_1[n]]^T$, $\alpha = 0.5$, and the cross-correlation vector and autocorrelation matrix are given by

$$\mathbf{r}_{xd} = \begin{bmatrix}1\\1\end{bmatrix}, \mathbf{R}_{xx} = \begin{bmatrix}2 & 1\\1 & 2\end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\1 & 1\end{bmatrix}\begin{bmatrix}3 & 0\\0 & 1\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\-1 & 1\end{bmatrix}.$$

Sketch the error surface (contour line) as a function of $h_0[n]$ and $h_1[n]$. Indicate clearly the minimum point and its value, the principal axis, and the direction of the contour lines

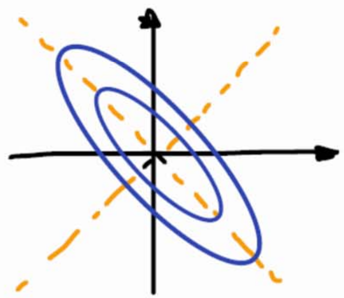
g)

$$\underline{\underline{R}}_x = \underline{\underline{U}} \cdot \underline{\underline{\Lambda}}\, \underline{\underline{V}}^T$$

$$\underline{\underline{R}}_x + \alpha \underline{\underline{I}} = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\1 & 1\end{bmatrix}\begin{bmatrix}7/2 & 0\\0 & 3/2\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\-1 & 1\end{bmatrix}$$

MAX EIGENVALUE : $\frac{7}{2} \longrightarrow \begin{bmatrix}1\\1\end{bmatrix}$

MIN EIGEN VALUE : $\frac{3}{2} \longrightarrow \begin{bmatrix}-1\\1\end{bmatrix}$

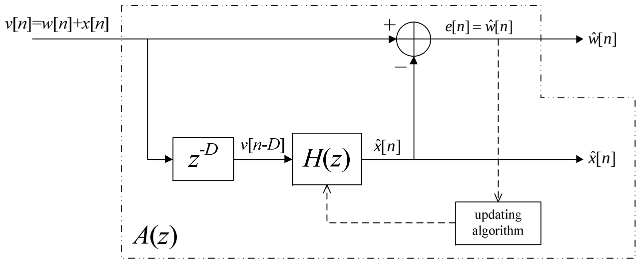


Examples of Optimal & Adaptive Filtering

1. **Affine predictor**
   - Comparison between linear and affine prediction for non-zero mean signals
2. **Wiener-Hopf solution for highly correlated data**
   - Avoiding the use of close to singular matrices
3. **Short term / Long term correlation**
   - Embedding a signal into noise

Short term / Long term correlation (I)

- The discrete process $v[n]$ has two additive and uncorrelated components which must be separated: a broad-band component $w[n]$ and a narrow-band component $x[n]$. As the band corresponding to $x[n]$ is unknown and can change along time, it is necessary to implement an adaptive filter:



It is well known that the autocorrelation of a broad-band signal has a lower effective length than the one of a narrow-band signal. Using this property, the scheme of the figure above allows to separate both signals provided that the delay D is large enough for $w[n]$ and $w[n-D]$ to be considered as uncorrelated samples, but small enough for $x[n]$ and $x[n-D]$ to be correlated. In practice, we will use the smallest value of D that meets this constraint. Consider all signals to be real and zero-mean. If $H(z)$ is a FIR filter with length M:

a) Show that the minimization of $E[e[n]^2]$ is equivalent to the minimization of $E[(x[n]-\hat{x}[n])^2]$, so that the scheme of the figure is able to separate the two components $w[n]$ and $x[n]$.

a) WE HAVE TO PROVE THAT:

$$\min_{\underline{h}} E\{e^2[n]\} = \min_{\underline{h}} E\{(x[n]-\hat{x}[n])^2\}$$

Short term / Long term correlation (II)

3.3

$$E\{e^2[n]\} = \left[e[n] = v[n] - \hat{x}[n] = w[n] - x[n] - \hat{x}[n] \right] =$$

$$= E\{(w[n] - x[n] - \hat{x}[n])^2\} = E\{(w[n] - (x[n] - \hat{x}[n]))^2\} =$$

$$= E\{w^2[n]\} - 2\, E\{w[n](x[n] - \hat{x}[n])\} + E\{(x[n] - \hat{x}[n])^2\}$$

$$E\{w[n](x[n] - \hat{x}[n])\} = E\{w[n]\,x[n]\} - E\{w[n]\,\hat{x}[n]\} =$$

$$= \left[E\{w[n+l]\,x[n]\} = \emptyset \right] = -E\{w[n]\,\hat{x}[n]\} = \left[\hat{x}[n] = \underline{h}^T \cdot \underline{v}[n-D] \right] =$$

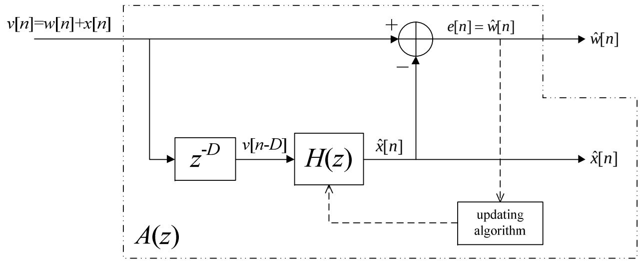
$$= -E\{w[n] \cdot \underline{h}^T \underline{v}[n-D]\} = \left[v[n] = w[n] + x[n] \right] =$$

$$= E\{w[n]\; \underline{h}^T (\underline{w}[n-D] + \underline{x}[n-D])\} = \underline{h}^T E\{w[n]\, \underline{w}[n-D]\} +$$

$$+ \underline{h}^T E\{w[n]\, \underline{x}[n-D]\} = \left[E\{w[n]\, w[n-D]\} = \emptyset \right] = \emptyset$$

Short term / Long term correlation (III)

a) Show that the minimization of $E[e[n]^2]$ is equivalent to the minimization of $E[(x[n]-\hat{x}[n])^2]$, so that the scheme of the figure is able to separate the two components $w[n]$ and $x[n]$:



$$\min_{h} E\{e^2[n]\} = \min_{h} E\{(x[n]-\hat{x}[n])^2\}$$

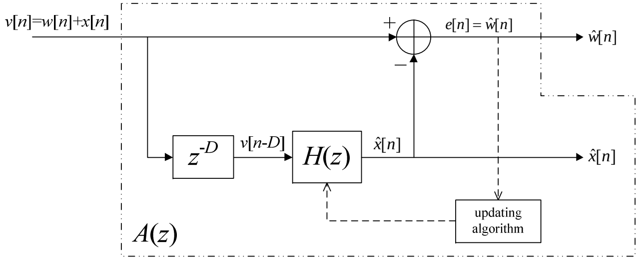
$$\min_{h} E\{e^2[n]\} = \min_{h} \left[E\{w^2[n]\} + E\{(x[n]-\hat{x}[n])^2\}\right]$$

AS $E\{w^2[n]\} = r_w[0]$, IT DOES NOT DEPEND ON h

$$\min_{h} E\{e^2[n]\} = \min_{h} E\{(x[n]-\hat{x}[n])^2\}$$

Short term / Long term correlation (IV)

b) Find the set of equations required to compute the coefficients of the filter $H(z)$.



b) FOLLOWING THE USUAL NOTATION :

$\Rightarrow$ OBSERVATIONS : $v[n-D] = w[n-D] + x[n-D]$

$\Rightarrow$ REFERENCE : $v[n] = w[n] + x[n]$

THEREFORE : $e[n] = v[n] - \hat{x}[n] = v[n] - \underline{h}^T [w[n-D] + x[n-D]]$

$E\{\underline{x}[n] e[n]\} = \underline{\emptyset} \Rightarrow E\{\underline{v}[n-D](v[n] - \hat{x}[n])\} = \underline{\emptyset}$

Short term / Long term correlation (V)

$$\underline{\underline{R}}_x \cdot \underline{h}_{opt} = \underline{r}_{xd} \Rightarrow E\{\underline{v}[n-D]\,\underline{v}^T[n-D]\}\,\underline{h}_{opt} = E\{\underline{v}[n-D]\,v[n]\}$$

$$E\{\underline{v}[n-D]\,\underline{v}^T[n-D]\} = E\{(\underline{w}[n-D] + \underline{x}[n-D])(\underline{w}[n-D] + \underline{x}[n-D])^T\} =$$

$$= E\{\underline{w}[n-D]\,\underline{w}^T[n-D]\} + E\{\underline{w}[n-D]\,\underline{x}^T[n-D]\} + E\{\underline{x}[n-D]\,\underline{w}^T[n-D]\} +$$

$$+ E\{\underline{x}[n-D]\,\underline{x}^T[n-D]\} = \underline{\underline{R}}_w + \underline{\underline{R}}_x$$

$$E\{\underline{v}[n-D]\,v[n]\} = E\{(\underline{w}[n-D] + \underline{x}[n-D])(w[n] + x[n])\} =$$

$$= E\{\underline{w}[n-D]\,w[n]\} + E\{\underline{w}[n-D]\,x[n]\} + E\{\underline{x}[n-D]\,w[n]\} +$$

$$+ E\{\underline{x}[n-D]\,x[n]\} = \underline{r}_x[D]$$

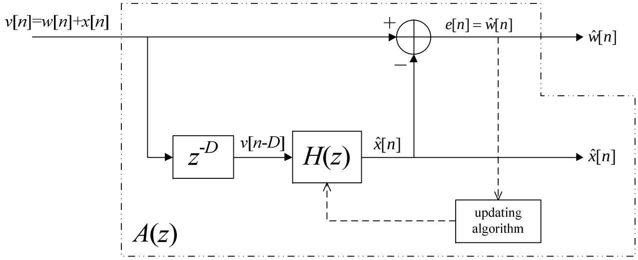
THEREFORE $\underline{\underline{R}}_x \cdot \underline{h}_{opt} = \underline{r}_{xd} \Rightarrow \left[\underline{\underline{R}}_w + \underline{\underline{R}}_x\right]\underline{h}_{opt} = \underline{r}_x[D]$

Short term / Long term correlation (VI)

3.3

In the following, we will assume that $x[n]$ is a sinusoid with frequency Ω_0 and $w[n]$ is white noise with variance σ^2.

3. Choose the proper values for D and M, and write the coefficients of the global filter encompassed within the dashed line in the figure, $A(z)$, in terms of the coefficients of the filter $H(z)$. Note that $A(z)$ is a prediction error filter, and give its impulse response.



c) $x[n] = A\cos\Omega_0 n$; $w[n] \longrightarrow r_w[\ell] = \sigma_w^2\,\delta[\ell]$

- THE MINIMUM DELAY IS $D=1$ SINCE $E\{w[n]w[n-1]\} = \emptyset$

- THE MINIMUM M VALUE IS TWO, SINCE WE HAVE TO FIX TWO PARAMETERS (A, Ω_0)

$$A(z) = \frac{E(z)}{V(z)} = 1 - H(z)z^{-D} = 1 - h_0 z^{-1} - h_1 z^{-2}$$

Short term / Long term correlation (VII)

4. Write the expression of the LMS equation that updates the filter coefficients.

d) $\underline{h}^{n+1} = \underline{h}^{n} + \mu \, \underline{x}[n] \, e[n]$ $\qquad \underline{x}[n] = \begin{bmatrix} v[n-1] \\ v[n-2] \end{bmatrix}$

$h_0^{n+1} = h_0^{n} + \mu \, v[n-1] \left[v[n] - \hat{x}[n] \right]$

$h_1^{n+1} = h_1^{n} + \mu \, v[n-2] \left[v[n] - \hat{x}[n] \right]$

Short term / Long term correlation (VIII)

3.3

5. If the power of the sinusoid is constant, reason how the speed of convergence and the misadjustment error change when the noise variance increases.

e) $v[n] = x[n] + w[n] \Rightarrow$ INC. $\Rightarrow \underline{\underline{R}}_v = \underline{\underline{R}}_x + \underline{\underline{R}}_w$

WITH $\underline{\underline{R}}_x = \begin{bmatrix} r_x[0] & r_x[1] \\ r_x[1] & r_x[0] \end{bmatrix}$ $\quad \underline{\underline{R}}_w = \sigma^2 \underline{\underline{I}} = \sigma^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

THE EIGENVALUES RESULT FROM $|\underline{\underline{R}}_x - \lambda^x \underline{\underline{I}}| = \emptyset$

AS $\underline{\underline{R}}_v = \underline{\underline{R}}_x + \sigma^2 \underline{\underline{I}} \Rightarrow |\underline{\underline{R}}_v - \sigma^2 \underline{\underline{I}} - \lambda^x \underline{\underline{I}}| = \emptyset$

$|\underline{\underline{R}}_v - (\sigma^2 + \lambda^x)\underline{\underline{I}}| = \emptyset \Rightarrow \lambda^v = \lambda^x + \sigma^2$

$\rightarrow N_{ITER} \propto \dfrac{\lambda_{MAX}}{\lambda_{MIN}} = \dfrac{\lambda^v_{MAX}}{\lambda^v_{MIN}} = \dfrac{\lambda_{MAX} + \sigma^2}{\lambda_{MIN} + \sigma^2} \Rightarrow \sigma \uparrow \quad N_{ITER} \downarrow$

$\rightarrow D \simeq \dfrac{\mu}{2} N r_x(\emptyset) = \dfrac{\mu}{2} N r_v(\emptyset) = \dfrac{\mu}{2} N (r_x(\emptyset) + \sigma^2) \Rightarrow \sigma \uparrow \quad D \uparrow$