5/3/21

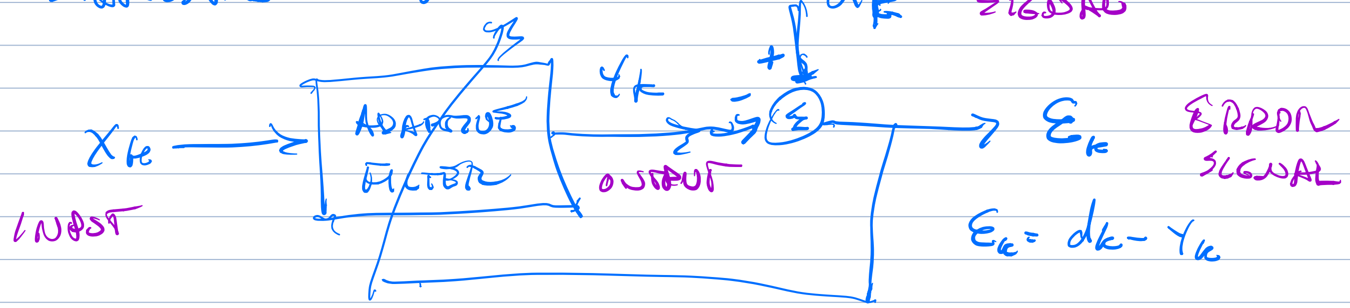
INTRO TO ADAPTIVE FILTERING

(STEARNS CHAP, LIM + OPPENHEIM NOTES, ADSP COURSE)

"NORMAL NOTATION" $x[n]$

ADAPTIVE FILTERING NOTATION X_k

STANDARD CONFIGURATION:



$\varepsilon_k = d_k - Y_k$

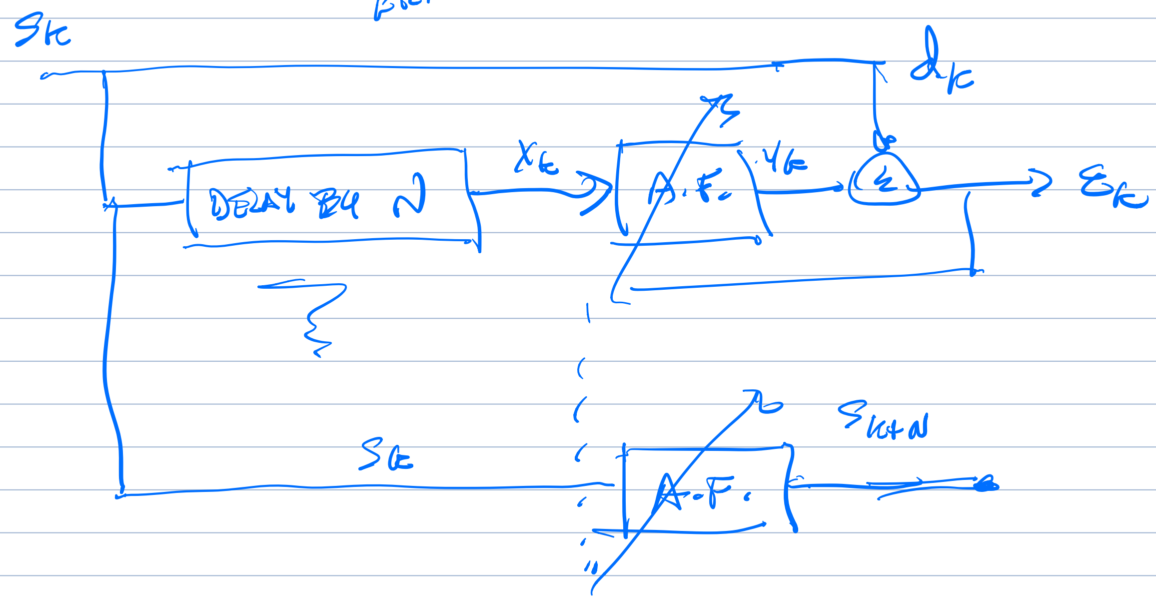
GOAL: TO MANIPULATE FILTER COEFFICIENTS TO MINIMIZE $E[\varepsilon_k^2]$

MEAN SQUARE ERROR → MSE

MIN. MEAN SQ. ESTIMATION → MMSE

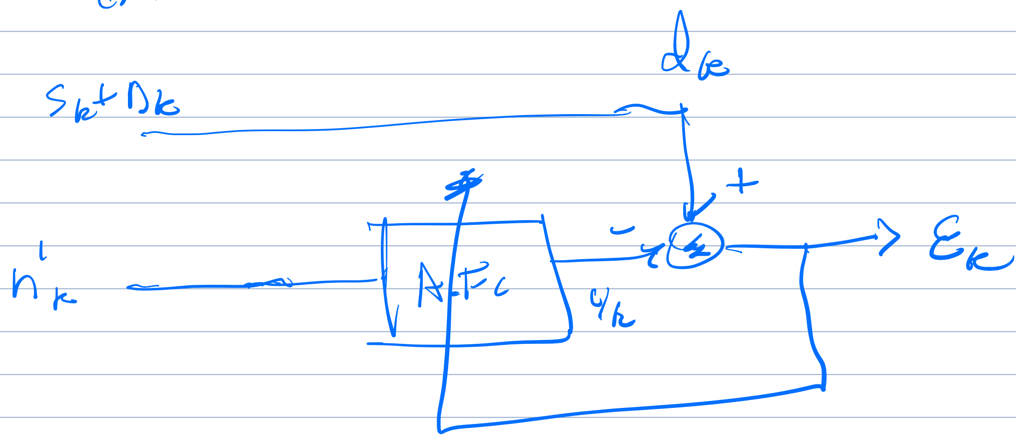
EXAMPLES

EX 1



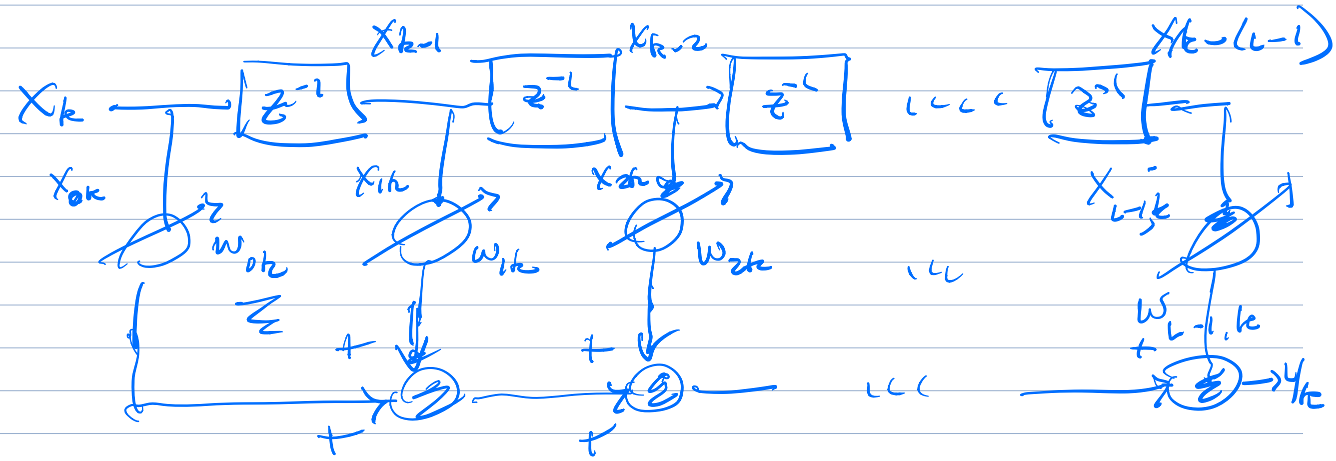
$X_k = s_{k-N}$

ADAPTIVE NOISE CANCELLATION

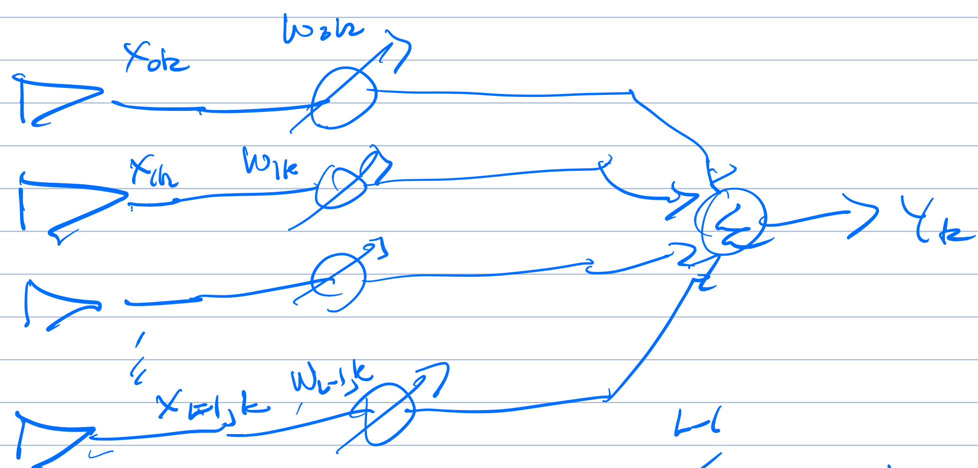


THE ADAPTIVE FILTER → ADAPTIVE LINEAR COMBINER:

EX 1: FIR FILTER OR TAPPED DELAY LINE (TDL)



EX. 2: SENSOR ARRAY



$$Y_k = \sum_{\ell=0}^{L-1} x_{\ell k} w_{\ell k} \quad \text{OR} \quad \sum_{\ell=0}^{L-1} x_{k-\ell} w_{\ell k}$$

SENSORS TDL

DEFINE $\underline{X}_k = \begin{bmatrix} x_{0k} \\ x_{1k} \\ x_{2k} \\ \vdots \\ x_{L-1,k} \end{bmatrix}$ $\underline{W}_k = \begin{bmatrix} w_{0k} \\ w_{1k} \\ \vdots \\ x_{L-1,k} \end{bmatrix}$

FOR TDL

$x_{lk} = x_{k-l}$

NOTE THAT $y_k = \underline{X}_k^T \underline{W}_k = \underline{W}_k^T \underline{X}_k$

$$\varepsilon_k = d_k - y_k = d_k - \underline{x}_k^T \underline{w}_k = d_k - \underline{W}_k \underline{X}_k$$

$$\varepsilon_k^2 = d_k^2 + \underline{w}_k^T \underline{x}_k \cdot \underline{x}_k^T \underline{w}_k - 2 d_k \underline{x}_k^T \underline{w}_k$$

y_k

$$\xi^2 = E[\varepsilon_k^2] = E[d_k^2] + \underline{w}_k^T E[\underline{x}_k \underline{x}_k^T] \underline{w}_k - 2E[d_k \underline{x}_k^T] \underline{w}_k$$

Let $\underline{\underline{R}} = E[\underline{x}_k \underline{x}_k^T]$, $\underline{P} = E[d_k \underline{x}_k]$

$L \times L$

$$\xi^2 = E[\varepsilon_k^2] = E[d_k^2] + \underline{w}^T \underline{\underline{R}}\, \underline{w} - 2 \underline{P}^T \underline{w}$$

COMMENTS:

1. ξ^2 IS A QUADRATIC FUNCTION of $\underline{w}_k$
2. IF $d_k, \underline{x}_k$ ARE STATIONARY, THEN SO ARE $\underline{\underline{R}}, \underline{P}$

GIVEN ξ^2, FIND $\underline{w}_k$ THAT MINIMIZES

$$\xi^2 = E[\varepsilon_k^2]$$

$$\xi^2 = E\{\varepsilon_k^2\} = E[d_k^2] + \underline{w}^T \underline{\underline{R}}\, \underline{w} - 2\underline{P}^T \underline{w}_k$$

Let $$\underline{\nabla} = \frac{d\xi^2}{d\underline{w}} = \begin{bmatrix} \frac{d\xi^2}{dw_0} \\ \frac{d\xi^2}{dw_1} \\ \vdots \\ \frac{d\xi^2}{dw_{L-1}} \end{bmatrix}$$

$$\underline{\nabla} = 2\underline{\underline{R}}\,\underline{w} - 2\underline{P} = \underline{0}$$

$$\cancel{2}\underline{\underline{R}}\,\underline{w} = \cancel{2}\underline{P}$$ WIENER-HOPF EQ.

$$\underline{w}_{OPT} = \boxed{\underline{w}^* = \underline{\underline{R}}^{-1}\underline{P}}$$

FINDING OPTIMAL $\underline{w}$ VIA GRADIENT DESCENT

Let $$\underline{w}_{k+1} = \underline{w}_k - \mu \underline{\nabla}_k$$

$$\underline{w}_{k+1} = \underline{w}_k - \mu\left(2\underline{\underline{R}}\,\underline{w} - 2\underline{P}\right)$$

$$= \underline{w}_k - 2\mu\underline{\underline{R}}\,\underline{w} + 2\mu\underline{P}$$

$$\underline{w}_{k+1} = \underline{w}_k - 2\mu E[\underline{x}_k \underline{x}_k^T]\,\underline{w}_k + 2\mu E[d_k \underline{x}_k]$$

PROBLEM: WE DON'T KNOW $\underline{\underline{R}} = E[\underline{x}\,\underline{x}^T]$, $\underline{P} = E[d_k \underline{x}_k]$

SOLUTION: REPLACE $\underline{\underline{R}} = E[\underline{x}\,\underline{x}^T]$ BY $\underline{x}_k \underline{x}_k^T$

$\underline{P} = E[d_k \underline{x}_k]$ BY $d_k \underline{x}_k$

STOCHASTIC GRADIENT DESCENT

$$\underline{w}_{k+1} = \underline{w}_k - 2\mu\, \underline{x}_k \underbrace{\underline{x}_k^T \underline{w}_k}_{y_k} + 2\mu\, d_k \underline{x}_k$$

$$\underline{w}_{k+1} = \underline{w}_k - 2\mu\, \underline{x}_k \underbrace{(y_k - d_k)}_{-\varepsilon_k}$$

$$\boxed{\underline{w}_{k+1} = \underline{w}_k + 2\mu\, \underline{x}_k \varepsilon_k}$$

THE LMS ALGORITHM !

LEAST MEAN SQUARES

(BERNARD WIDROW)