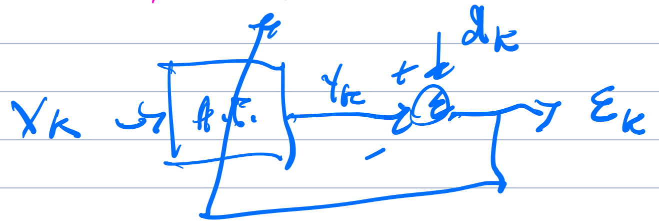
11/11/24

Performance of Adaptive Filters; The RLS Algorithm

(LO 5.4–5.6; Notes 8.5–8.7)

Reminder: Q2 next week



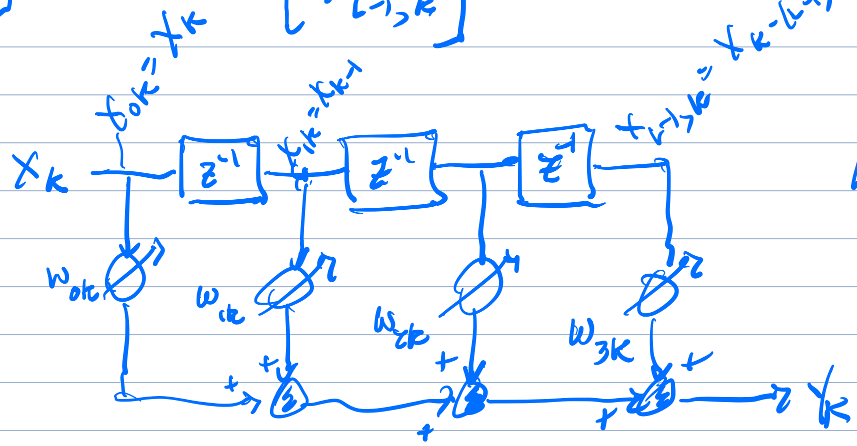
$y_k = \underline{w}_k^T \underline{x}_k = \underline{x}_k^T \underline{w}_k$

Find $\{\underline{w}_k\}$ to min $E[\varepsilon_k^2]$

$$\underline{w}_k = \begin{bmatrix} w_{0k} \\ w_{1k} \\ w_{2k} \\ \vdots \\ w_{L-1,k} \end{bmatrix} \qquad \underline{x}_k = \begin{bmatrix} x_{0k} \\ x_{1k} \\ x_{2k} \\ \vdots \\ x_{L-1,k} \end{bmatrix}$$

FIR filter

$x_{ik} = x_{k-i}$



$L = 4$

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

where $\underline{\underline{R}} = E\{\underline{x}_k \underline{x}_k^T\}$ $L \times L$

$\underline{P} = E[d_k \underline{x}_k]$ $L \times 1$

Assume $\underline{\underline{R}}$, $\underline{P}$ are stationary, known

$$\frac{\partial \xi^2}{\partial w_k} = \frac{\partial \nabla_{\xi^2}}{\partial w_k} = \cancel{2}\,\underline{\underline{R}}\,\underline{w} - \cancel{2}\,\underline{P} = 0$$

$$\underline{\underline{R}}\,\underline{w} = \underline{P}$$

$$\cancel{\underline{\underline{R}}^{-1}\underline{\underline{R}}}\,\underline{w} = \underline{\underline{R}}^{-1}\underline{P}$$

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

WIENER-HOPF

NEWTON SOLUTION

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

$$\frac{1}{2}\underline{\underline{R}}^{-1}\underline{\nabla} = \frac{1}{\cancel{2}}\cancel{2}\,\cancel{\underline{\underline{R}}^{-1}\underline{\underline{R}}}\,\underline{w} - \underbrace{\cancel{\tfrac{2}{2}}\underline{\underline{R}}^{-1}\underline{P}}_{\underline{w}^{*}}$$

$$\frac{1}{2}\underline{\underline{R}}^{-1}\underline{\nabla} = \underline{w} - \underline{w}^{*}$$

$$\boxed{\underline{w}^{*} = \underline{w} - \frac{1}{2}\underline{\underline{R}}^{-1}\underline{\nabla}}$$

GRADIENT DESCENT

$$\underline{w}_{k+1} = \underline{w}_k - \mu\,\frac{1}{2}\underline{\underline{R}}^{-1}\underline{\nabla} \qquad 0 < \mu < 1$$

PROBLEMS: 1. DON'T KNOW $\underline{\underline{R}}$

2. DON'T KNOW $\underline{\nabla}$

LMS ALGORITHM

1. IGNORE $\underline{\underline{R}}^{-1}$

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

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

"STOCHASTIC GRADIENT DESCENT"

$$W_{k+1} = W_k - \mu \frac{1}{2} \cancel{R}^{-1} \nabla$$

$$W_{k+1} = W_k - \mu \frac{1}{\cancel{2}} \underbrace{\left(\cancel{2} R W - \cancel{2} P \right)}_{\nabla_{\varepsilon^2}}$$

$$W_{k+1} = W_k - \mu \left(X_k \underbrace{X_k^T W_k}_{y_k} - d_k X_k \right)$$

$$W_{k+1} = W_k - \mu X_k \left(\underbrace{y_k - d_k}_{-\varepsilon_k} \right)$$

$$\boxed{W_{k+1} = W_k + \mu X_k \varepsilon_k}$$ LMS

PERFORMANCE of GRADIENT DESCENT

$$W_{k+1} = W_k - \mu \cancel{\tfrac{1}{2}} R^{-1} \nabla$$

$$W_{k+1} = W_k - \mu \frac{1}{\cancel{2}} R^{-1} \left(2 R W_k - 2P \right)$$

$$W_{k+1} = W_k - \mu\, 2 W_k + \mu 2 \underbrace{R^{-1} P}_{W^*}$$

$$\boxed{W_{k+1} = (1-2\mu) W_k + 2\mu W^*}$$

ASSUME ARBITRARY W_0, ITERATE

$$W_1 = (1-2\mu) W_0 + 2\mu W^*$$

$$W_2 = (1-2\mu) W_1 + 2\mu W^*$$

$$= (1-2\mu)\left[(1-2\mu) W_0 + 2\mu W^* \right] + 2\mu W^*$$

$$= (1-2\mu)^2 W_0 + 2\mu (1-2\mu) W^* + 2\mu W^*$$

$$W_3 = (1-2\mu) W_2 + 2\mu W^*$$

$$W_3 = (1-2\mu)\left[(1-2\mu)^2 W_0 + 2\mu (1-2\mu) W^* + 2\mu W^* \right] + 2\mu W^*$$

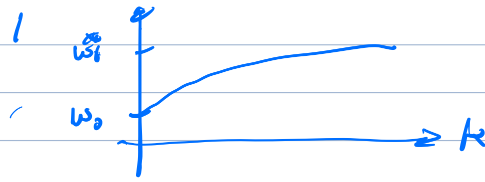
$$w_3 = (1-2\mu)^3 w_0 + (1-2\mu)^2 2\mu w^* + (1-2\mu)2\mu w^* + 2\mu w^*$$

$$w_k = (1-2\mu)^k w_0 + 2\mu w^* \sum_{\ell=0}^{k-1} (1-2\mu)^\ell$$

$$= (1-2\mu)^k w_0 + \cancel{2\mu}\, w^* \frac{1-(1-2\mu)^k}{\underbrace{\cancel{1-(1-2\mu)}}_{2\mu}}$$

$$w_k = (1-2\mu)^k w_0 + w^*\left(1-(1-2\mu)^k\right)$$

IF $|1-2\mu| < 1$



FOR $1-2\mu$ POSITIVE

$$(1-2\mu) < 1 \Rightarrow \mu < 1/2$$

FOR $1-2\mu$ NEGATIVE

$$-(1-2\mu) < 1 \Rightarrow \mu > 0 \Rightarrow 0 < \mu < \frac{1}{2}$$

CONVERGENCE OF STOCHASTIC GRADIENT DESCENT:

THE RLS ALGORITHM

INCREMENTAL NEWTON SOLUTION

NEWTON: $w_{k+1} = w_k - \mu R^{-1} \nabla$

LMS: $w_{k+1} = w_k - \mu \hat{\nabla}$

RLS: $w_{k+1} = w_k - \mu \hat{R}^{-1} \hat{\nabla}$

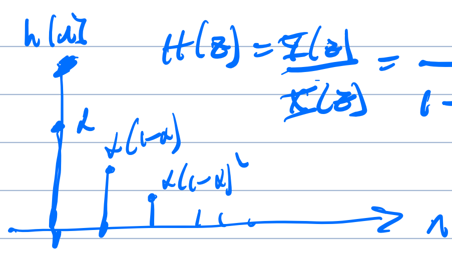
$$R = E[x_k x_k^T] = \langle x_k x_k^T \rangle \text{ IF ERGODIC}$$

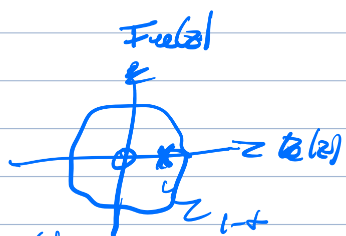
$$x_k \to \boxed{h_k} \to y_k$$

$$y[n] = \alpha x[n] + (1-\alpha) y[n-1] \qquad 0 < \alpha < 1$$

$$Y(z) = \alpha X(z) + (1-\alpha) z^{-1} Y(z)$$

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\alpha}{1-(1-\alpha)z^{-1}} = \frac{z\alpha}{z-(1-\alpha)}$$





$$h[n] = \alpha(1-\alpha)^n u[n]$$

FOR ESTIMATING $\underline{R}$: $\quad \underline{x}_k \underline{x}_k^T \longrightarrow \boxed{LSI} \longrightarrow \hat{\underline{R}}_k$

FROM $y[n] = \alpha x[n] + (1-\alpha)\, y[n-1]$

$$\hat{\underline{R}}_k = \alpha\, \underline{x}_k \underline{x}_k^T + (1-\alpha)\, \hat{R}_{k-1}$$

$$\hat{\underline{R}}_k = (1-\alpha)\, \hat{R}_{k-1} + \alpha\, \underline{x}_k \underline{x}_k^T$$

PRE-MULT BY $\hat{\underline{R}}_k^{-1}$, POST-MULT BY $\hat{\underline{R}}_{k-1}$

$$\cancel{\hat{\underline{R}}_k^{-1}}\, \cancel{\hat{\underline{R}}_k}\, \hat{\underline{R}}_{k-1} = (1-\alpha)\, \hat{\underline{R}}_k^{-1} \cancel{\hat{\underline{R}}_{k-1}}\, \cancel{\hat{\underline{R}}_{k-1}^{-1}} + \alpha\, \underline{R}_k^{-1} \underline{x}_k \underline{x}_k^T \hat{\underline{R}}_{k-1}^{-1}$$

(A) $$\hat{\underline{R}}_{k-1}^{-1} = (1-\alpha)\, \hat{\underline{R}}_k^{-1} + \alpha\, \hat{\underline{R}}_k^{-1} \underline{x}_k \underline{x}_k^T \hat{\underline{R}}_{k-1}^{-1}$$

POST MULT BY $\underline{x}_k$:

$$\underbrace{\hat{\underline{R}}_k^{-1} \underline{x}_k}_{\underline{g}_k} = (1-\alpha)\, \hat{\underline{R}}_k^{-1} \underline{x}_k + \alpha\, \underbrace{\hat{\underline{R}}_k^{-1} \underline{x}_k}_{\underline{g}} \underline{x}_k^T \underbrace{\underline{R}_k^{-1} \underline{x}_k}_{\underline{g}_k}$$

DEFINE $\underline{g}_k \triangleq \hat{\underline{R}}_{k-1}^{-1} \underline{x}_k$

$$\underline{g}_k = \underline{R}_k^{-1} \underline{x}_k \left[(1-\alpha) + \alpha\, \underline{x}_k^T \underline{g}_k\right]$$

DIVIDE BY $(1-\alpha) + \alpha\, \underline{x}_k^T \underline{g}_k$, POST MULT BY $\underline{g}_k^T = \underline{x}_k^T \underline{R}_{k-1}^{-1}$

(B) $$\frac{\underline{g}_k \underline{g}_k^T}{(1-\alpha) + \alpha\, \underline{x}_k^T \underline{g}_k} = \hat{\underline{R}}_k^{-1} \underline{x}_k \underline{x}_k^T \hat{\underline{R}}_{k-1}^{-1}$$

(A) $$\hat{\underline{R}}_{k-1}^{-1} = (1-\alpha)\, \hat{\underline{R}}_k^{-1} + \alpha\, \hat{\underline{R}}_k^{-1} \underline{x}_k \underline{x}_k^T \hat{\underline{R}}_{k-1}^{-1}$$

$$\hat{\underline{R}}_{k-1}^{-1} = (1-\alpha)\, \hat{\underline{R}}_k^{-1} + \alpha \frac{\underline{g}_k \underline{g}_k^T}{(1-\alpha) + \alpha\, \underline{x}_k^T \underline{g}_k}$$

$$\boxed{\underline{R}_k^{-1} = \frac{1}{1-\alpha}\left(\hat{\underline{R}}_{k-1}^{-1} + \frac{\alpha\, \underline{g}_k \underline{g}_k^T}{\left((1-\alpha) + \alpha\, \underline{x}_k^T \underline{g}_k\right)}\right)}$$

THE RLS PROCEDURE

1. Let $\hat{\underline{\underline{R}}}_0^{-1} = \underline{\underline{R}}^{-1}$ or $\frac{1}{\sigma_x^2}\underline{\underline{I}}$

2. Let $\hat{\underline{S}}_k = \hat{\underline{\underline{R}}}_{k-1}^{-1}\underline{X}_k$

3. UPDATE $\hat{R}_k^{-1} = \frac{1}{1-\alpha}\left[\hat{\underline{\underline{R}}}_{k-1} - \frac{\alpha\,\hat{\underline{S}}_k\hat{\underline{S}}_k^T}{1-\alpha+\alpha\underline{X}_k^T\hat{\underline{S}}_k}\right]$

4. $\varepsilon_k = d_k - \underline{X}_k^T\underline{W}_k$

5. UPDATE $\underline{W}_{k+1} = \underline{W}_k + 2\mu\,\varepsilon_k\,\hat{\underline{\underline{R}}}_k^{-1}\,\underline{X}_k$