EE648 (CC761-M) DSP II
Session 2 (Live: 1/14/99)

<u>Outline</u>:

- Review MMSE Criterion for Adaptive Filters – Sect. 12.2.1 1st Ed. of P+M
- Widrow LMS Algorithm – Sect. 12.2.2, 1st Ed. of P+M
- Convergence analysis of LMS
- Sect. 12.2.3, 1st Ed. of P+M

- recall: $E\{e^2[n]\}$

$$= E\{d^2[n]\} - \underline{r}_{dx}^T \underline{R}_{xx}^{-1} \underline{r}_{dx} + (\underline{R}_{xx}\underline{h}_M - \underline{r}_{dx})^T \underline{R}_{xx}^{-1} (\underline{R}_{xx}\underline{h}_M - \underline{r}_{dx})$$

- optimum $\underline{h}_M$ satisfies

$$\underline{R}_{xx}\underline{h}_M - \underline{r}_{dx} = \underline{0} \Rightarrow \underline{h}_M^{opt} = \underline{R}_{xx}^{-1}\underline{r}_{dx}$$

- Minimum value of $E\{e^2[n]\}$:

$$= E\{d^2[n]\} - \underline{r}_{dx}^T (\underline{R}_{xx}^{-1}\underline{r}_{dx})$$

$$= E\{d^2[n]\} - \underline{r}_{dx}^T \underline{h}_M^{opt}$$

- in practice, don't know $\underline{R}_{xx}$ or $\underline{r}_{dx}$
- must estimate these quantities
- also: statistics of $x[n]$ and $d[n]$ may vary "slowly" with time
- adaptive approach:
- employ gradient-based search to iterate towards the filter coeff. vector minimizing $E\{e^2[n]\}$
- error surface is a hyperparaboloid in M-dimensional space

- search for "bottom of bowl"
- let gradient vector evolve with time in accordance with $\{x[n]\}$ and $\{d[n]\}$
- gradient operator:

$$\nabla_{\underline{h}_M} = \left[\frac{\partial}{\partial h[0]}, \frac{\partial}{\partial h[1]}, \cdots, \frac{\partial}{\partial h[M-1]}\right]^T$$

- $E\{e^2[n]\} = f(\underline{h}_M)$

$$= E\{d^2[n]\} - 2\underline{h}_M^T \underline{r}_{dx} + \underline{h}_M^T \underline{R}_{xx} \underline{h}_M$$

• easy to show:

$$\nabla_{\underline{h}_M}\left(\underline{h}_M^T \underline{r}_{dx}\right) = \underline{r}_{dx} \quad (M\times 1)$$

$$\nabla_{\underline{h}_M}\left(\underline{h}_M^T \underline{R}_{xx} \underline{h}_M\right) = 2 \underbrace{\underline{R}_{xx}}_{M\times M} \underbrace{\underline{h}_M}_{M\times 1}$$

• thus:

$$\nabla_{\underline{h}_M} f(\underline{h}_M) = -2\,\underline{r}_{dx} + 2\,\underline{R}_{xx}\underline{h}_M$$

$$= \underline{0} \Rightarrow \underline{R}_{xx}\underline{h}_M = \underline{r}_{dx}$$

$$\Rightarrow \underline{h}_M^{opt} = \underline{R}_{xx}^{-1}\,\underline{r}_{dx}$$

• consider using method of "steepest descent" (classical gradient search)

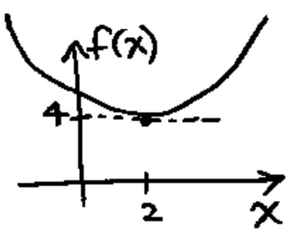
$$\underline{h}_M(\ell+1) = \underline{h}_M(\ell) - \frac{1}{2}\mu_\ell \nabla_{\underline{h}_M} f(\underline{h}_M^{(\ell)})$$

• each iteration: take "step" in direction of negative gradient

• μ_ℓ: step size $(0 < \mu_\ell < 1)$

• Simple example:

$$f(x) = x^2 - 4x + 8 = (x-2)^2 + 4$$



Min $f(x) = 4$ at $x = 2$

$$\nabla_x f(x) = \frac{d}{dx} f(x) = 2(x-2)$$

• gradient search:

$$x(\ell+1) = x(\ell) - \frac{1}{2}\mu_\ell\, 2(x^{(\ell)}-2)$$

$$= x(\ell) - \mu_\ell (x^{(\ell)}-2)$$

• if $x(\ell) < 2 \Rightarrow -\mu_\ell(x(\ell)-2) > 0$

$\Rightarrow$ take step in $+x$ direction

• if $x(\ell) > 2 \Rightarrow -\mu_\ell(x(\ell)-2) < 0$

$\Rightarrow$ take step in $-x$ direction

• Widrow's LMS Algorithm

• Key features:

• fix step size μ_ℓ

• approximate $\underline{r}_{dx}$ and $\underline{R}_{xx}$ by instantaneous estimates

$$\hat{\underline{r}}_{dx}[n] = d[n]\,\underline{x}[n]$$

$$\hat{\underline{R}}_{xx}[n] = \underline{x}[n]\,\underline{x}^T[n]$$

• Substitute these into the expression for the gradient

$$\nabla_{\underline{h}_M} E\{e^2[n]\} = -2\,\underline{r}_{dx} + 2\,\underline{R}_{xx}\,\underline{h}_M$$

$$\hat{\nabla}_{\underline{h}_M[n]} = -2\, d[n]\,\underline{x}[n] + 2\,\underline{x}[n]\underline{x}^T[n]\,\underline{h}_M[n]$$

$$= -2\left\{d[n] - \underline{x}^T[n]\underline{h}_M[n]\right\}\underline{x}[n]$$

$$= -2\left\{d[n] - \hat{d}[n]\right\}\underline{x}[n]$$

$$= -2\, e[n]\,\underline{x}[n]$$

$$\underline{h}_M[n+1] = \underline{h}_M[n] - \tfrac{1}{2}\mu\, \hat{\nabla}_{\underline{h}_M} E\{e^2[n]\}$$

$$= \underline{h}_M[n] + \mu\, e[n]\,\underline{x}[n]$$

- Convergence Analysis of LMS
- show LMS udpate converges to $\underline{h}_M^{opt} = \underline{R}_{xx}^{-1}\underline{r}_{dx}$ in the mean provided $\{x[n]\}$ is stationary and $0 < \mu < \frac{2}{\lambda_{max}}$
- where λ_{max} is largest eigenvalue of $\underline{R}_{xx}$

- Proof: define

$$\underline{c}[n] = \underline{\bar{h}}_M[n] - \underline{h}_M^{opt}$$

- where: $\underline{\bar{h}}_M[n] = E\{\underline{h}_M[n]\}$
- thus: $\underline{c}[n+1] = \underline{\bar{h}}_M[n+1] - \underline{h}_M^{opt}$
- take expected value of both sides of LMS update equation:

$$\underline{\bar{h}}_M[n+1] = \underline{\bar{h}}_M[n] + \mu\, E\{e[n]\,\underline{x}[n]\}$$

- Subtract $\underline{h}_M^{opt}$ from both sides

$$\underline{c}[n+1] = \underline{c}[n] + \mu\, E\{e[n]\,\underline{x}[n]\}$$

$$= \underline{c}[n] + \mu\, E\{d[n]\,\underline{x}[n]\} - \mu\, E\{\underline{x}[n]\,\underbrace{\underline{x}^T[n]\,\underline{h}_M[n]}_{\hat{d}[n]}\}$$

$$= \underline{c}[n] + \mu\, \underline{r}_{dx} - \mu\, E\{\underline{x}[n]\underline{x}^T[n]\underline{h}_M[n]\}$$

- asymptotically:

$$E\{\underline{x}[n]\underline{x}^T[n]\left(\underline{\bar{h}}_M[n] + \Delta\underline{h}_M[n]\right)\}$$

$$= \underline{R}_{xx}\,\underline{\bar{h}}_M[n] + E\{\underline{x}[n]\underline{x}^T[n]\,\Delta\underline{h}_M[n]\}$$

- assuming $\Delta \underline{h}_M[n] << 0$,

$$\underline{c}[n+1] = \underline{c}[n] + \mu\, \underline{r}_{dx} - \mu\, \underline{R}_{xx}\, \overline{\underline{h}}_M[n]$$

$$= \underline{c}[n] + \mu\, \underline{R}_{xx}\, \underline{R}_{xx}^{-1} \underline{r}_{dx} - \mu\, \underline{R}_{xx}\, \overline{\underline{h}}_M[n]$$

$$= \underline{c}[n] + \mu\, \underline{R}_{xx} \left\{ \underline{h}_M^{opt} - \overline{\underline{h}}_M[n] \right\}$$

$$= \underline{c}[n] - \mu\, R_{xx}\, \underline{c}[n]$$

$$\underline{c}[n+1] = \left\{ \underline{I} - \mu\, \underline{R}_{xx} \right\} \underline{c}[n]$$

- consider eigenvalue decomposition of $\underline{R}_{xx}$

$$\underline{R}_{xx} = \underline{U}\, \underline{\Lambda}\, \underline{U}^T$$

symmetric

- since $\underline{R}_{xx}$ is positive-definite +

$$\left.\underline{U}^T \underline{U} = \underline{I} = \underline{U}\,\underline{U}^T \right\} \text{eigenvectors are orthonormal}$$

$$\underline{c}[n+1] = \left\{ \underline{U}\,\underline{U}^T - \mu\, \underline{U}\, \underline{\Lambda}\, \underline{U}^T \right\} \underline{c}[n]$$

$$\underline{U}^T \underline{c}[n+1] = \left\{ \underline{I} - \mu\, \underline{\Lambda} \right\} \underline{U}^T \underline{c}[n]$$

- define $\underline{c}^o[n+1] = \underline{U}^T \underline{c}[n+1]$

$$\underline{c}^o[n+1] = \left\{ \underline{I} - \mu\, \underline{\Lambda} \right\} \underline{c}^o[n]$$

- component-wise:

$$c^o[k;n+1] = (1 - \mu \lambda_k)\, c^o[k;n] \qquad k = 1, \ldots, M$$

- recall:

$$h[n] = a\, h[n-1] \quad \left(h[n+1] = a\, h[n] \right)$$

soln: $h[n] = a^n\, h[0]$

- thus: $c^o[k;n] = (1 - \mu \lambda_k)^n c^o[k;0]$

$$k = 1, \ldots, M$$

- for convergence, require:

$$-1 < 1 - \mu \lambda_k < 1 \quad \text{for } k = 1, \ldots, M$$

$$0 < \mu < \frac{2}{\lambda_k}$$

- to insure convergence:

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

- in practice: $\lambda_{max} < \sum_{k=1}^{M} \lambda_k = \text{trace}\left\{ \underline{R}_{xx} \right\} = M r_{xx}[0]$

- $0 < \mu < \dfrac{2}{M r_{xx}[0]}$