Paper Number(s): **E3.08**
ISE3.1

IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE
UNIVERSITY OF LONDON

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2000

EEE/ISE PART III/IV: M.Eng., B.Eng. and ACGI

ADVANCED SIGNAL PROCESSING

Tuesday, 16 May 2000, 10:00 am

There are FIVE questions on this paper.

Answer ONE question from Section A, and TWO from Section B.

Use the same answer book for each section.

Time allowed: 3:00 hours

Corrected Copy

None

Examiners: Dr J.A. Chambers, Prof A.G. Constantinides

Special instructions for invigilators:	One main answer book only is needed on each desk (not one each for Sections A and B).
Information for candidates:	Write your answers for Sections A and B in the same answer book.

Section A

1.

The power spectral density, $P_x(e^{j2\pi f})$, of a real, zero mean, wide sense stationary discrete time random signal, x[n], is related to its autocorrelation sequence, $r_x(\tau)$, by

$$P_x(e^{j2\pi f}) = F[r_x(\tau)] \quad f \in (-0.5, 0.5]$$

where F[.] denotes the discrete Fourier transform.

(a) Verify and discuss the following properties of the power spectral density of x[n]

(i) $P_x(e^{j2\pi f}) = P_x^*(e^{j2\pi f})$

(ii) $P_x(e^{j2\pi f}) = P_x(e^{-j2\pi f})$

(iii) $P_x(z) = P_x^*(1/z^*)$

where $(.)^*$ denotes complex conjugate, and z is the complex variable in the z-transform.

(b) If y[n] is the output of a linear system with input x[n], transfer function $H(e^{j2\pi f})$, and $P_y(e^{j2\pi f}) = |H(e^{j2\pi f})|^2 P_x(e^{j2\pi f})$, show that

$$P_x(e^{j2\pi f}) \geq 0, \forall f.$$

(c) Calculate and sketch the autocorrelation sequences that correspond to the following expressions

(i) $P_x(e^{j2\pi f}) = 4 + 2\cos 2\pi f$

(ii) $P_x(e^{j2\pi f}) = \dfrac{2}{5 + 3\cos 2\pi f}$

(iii) $P_x(z) = \dfrac{-4z^2 + 10z - 4}{3z^2 + 10z + 3}$

2.

(a) List the conditions for a real discrete time random signal, x[n], to be wide sense stationary.

(b) The mean ergodic theorem states that a necessary and sufficient condition for x[n] to be ergodic in the mean is that its autocovariance sequence, $c_x(\tau)$, must satisfy

$$\lim_{N\to\infty} \frac{1}{N}\sum_{\tau=0}^{N-1} c_x(\tau) = 0.$$

Hence, or otherwise, determine whether the following discrete time random signals are wide sense stationary and mean ergodic

(i) $x[n] = \theta$, where θ is a random variable which has probability density $p_\Theta(\theta)$.

(ii) $x[n] = A\cos(2\pi n f_o + \phi)$ where A and f_o are constants and ϕ is a uniformly distributed random variable between -π and π.

(iii) x[n] is a Bernoulli discrete time random signal with Pr{x[n] = 1} = p and Pr{x[n] = -1} = 1-p.

Section B

3.

(a) Discuss the term BLUE estimator and the information that is required for its formulation given a real observation data set {x[0], x[1], ..., x[N-1]} whose joint probability density function is dependent upon an unknown $p \times 1$ parameter vector $\underline{\theta}$.

Suppose that the observation dataset satisfies the vector model

$$\underline{x} = H\underline{\theta} + \underline{w}$$

where $\underline{x}$ is an $N \times 1$ vector of data observations, H is a known $N \times p$ observation matrix, with $N > p$ and full column rank, and $\underline{w}$ is an $N \times 1$ vector of zero mean noise terms.

(b) Verify that the BLUE estimator is given by

$$\hat{\underline{\theta}} = (H^T C^{-1} H)^{-1} H^T C^{-1} \underline{x}$$

in which C is the observation vector covariance matrix and $(.)^T$ denotes vector transpose.

(c) By considering the affine transformation

$$\underline{\alpha} = B\underline{\theta} + \underline{b}$$

where B is a known $p \times p$ invertible matrix and $\underline{b}$ is a known $p \times 1$ vector, prove that the BLUE estimator commutes over linear transformations of $\underline{\theta}$.

4.

(a) Show in block diagram form how an adaptive filter can be employed to enhance the operation of a speech recognition system within an in-car hands-free mobile phone.

(b) Derive the least mean square (LMS) adaptive algorithm from the method of steepest descent which is based upon the minimization of the mean squared error

$$J = E\{e^2[n]\}$$

where $e[n] = d[n] - \underline{w}^T[n]\underline{x}[n]$, $d[n]$ is the desired response, $\underline{w}[n]$ is the $p \times 1$ parameter vector of the adaptive filter and $\underline{x}[n]$ is the input vector of the adaptive filter $[x[n], x[n-1], \ldots, x[n-p+1]]^T$.

(c) Calculate the theoretical minimum mean square error of the filter in (b) and explain whether the LMS algorithm can attain this performance.

(d) The robust mixed norm (RMN) adaptive algorithm minimizes the instantaneous cost function

$$J = \lambda e^2[n] + (1-\lambda)|e[n]|$$

where $\lambda \in [0,1]$ is a scalar mixing parameter.

(i) Show the parameter update equation for the RMN algorithm.

(ii) Discuss the advantages and disadvantages of the RMN algorithm as compared to the LMS algorithm.

(iii) Suggest a scheme for on-line selection of λ.

5.

(a) Discuss the difference between a block-based and a sequential estimator.

(b) State the orthogonality principle of least squares estimation given the real vector signal model $\underline{s}[n] = H\underline{\theta}$ for the $N \times 1$ vector of data observations, where H is a known $N \times p$ observation matrix, with $N > p$ and full column rank, and $\underline{\theta}$ is a $p \times 1$ parameter vector.

(c) Using the orthogonality principle, or otherwise, calculate the block-based least squares estimator for $\underline{\theta}$.

(d) Show that the minimum least squares error of the estimator in (c) can be written as

$$J_{LS} = \underline{x}^T (I - H(H^T H)^{-1} H^T)\underline{x}$$

(e) Convert the block-based estimator for $\underline{\theta}$ into a sequential least squares estimator.

1) $P_x(e^{j2\pi f}) = \sum_{\tau=-\infty}^{\infty} r_x(\tau) e^{-j2\pi f\tau}$ (Normalised f is continuous)

Key point, as $x[n]$ is real, zero mean, and WSS, $r_x(\tau) = r_x(-\tau)$

$r_x(\tau) = r_x^*(\tau) = r_x^*(-\tau)$

(i) Hence

(a) $P_x^*(e^{j2\pi f}) = \left(\sum_{\tau=-\infty}^{\infty} r_x(\tau) e^{-j2\pi f\tau}\right)^* = \sum_{\tau=-\infty}^{\infty} r_x^*(\tau) e^{j2\pi f\tau}\Big|_{\tau=-s}$

$= \sum_{s=-\infty}^{\infty} r_x^*(-s) e^{-j2\pi fs} = \sum_{s=-\infty}^{\infty} r_x(s) e^{-j2\pi fs} = P_x(e^{j2\pi f})$

$\Rightarrow P_x(e^{j2\pi f})$ is real.

(b) $P_x(e^{-j2\pi f}) = \left(\sum_{\tau=-\infty}^{\infty} r_x(\tau) e^{j2\pi f\tau}\right)\Big|_{\tau=-s} = \sum_{s=-\infty}^{\infty} r_x(-s) e^{-j2\pi fs} = \sum_{s=-\infty}^{\infty} r_x(s) e^{-j2\pi fs}$

$= P_x(e^{j2\pi f})$

$\Rightarrow P_x(e^{j2\pi f})$ is symmetric.

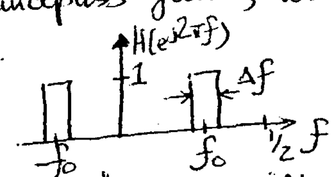
(c) $P_x(z) = P_x(e^{j2\pi f})\Big|_{z=e^{j2\pi f}}$

$= P_x(z) = \sum_{\tau=-\infty}^{\infty} r_x(\tau) z^{-\tau}$ (NB. Bi-lateral z-transform, note ROC) (8)

$= P_x^*\left(\frac{1}{z^*}\right) = \left(\sum_{\tau=-\infty}^{\infty} r_x(\tau)\left(\frac{1}{z^*}\right)^{-\tau}\right)^* = \sum_{\tau=-\infty}^{\infty} r_x^*(\tau) z^{\tau}\Big|_{\tau=-s} = \sum_{s=-\infty}^{\infty} r_x(s) z^{-s} = P_x(z)$

$\Rightarrow$ If $P_x(z)$ a rational function, poles and zeros lie in conjugate reciprocal pairs, leads to spectral factorisation.

(ii) Consider $H(e^{j2\pi f})$ to be an ideal narrow-bandpass filter, with arbitrary centre frequency, f_0, and bandwidth, Δf, i.e.



if $x[n]$ is filtered by $H(e^{j2\pi f})$, then the output $y[n]$ will have psd $P_y(e^{j2\pi f}) = |H(e^{j2\pi f})|^2 P_x(e^{j2\pi f})$. Therefore, the average power within $y[n]$, $E\{y^2[n]\} = r_y(0) = \int_{-1/2}^{1/2} |H(e^{j2\pi f})|^2 P_x(e^{j2\pi f})\,df =$

$2\int_{f_0-\Delta f/2}^{f_0+\Delta f/2} P_x(e^{j2\pi f})\,df \cong 2\Delta f\, P_x(e^{j2\pi f_0})$, and by definition $E\{y^2[n]\} \geq 0$,

hence $P_x(e^{j2\pi f}) \geq 0 \;\forall f$ ∎

(8)

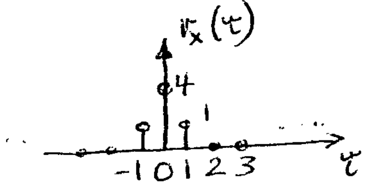
J.C.

1) Cont.

(iii)

a) $P_x(e^{j2\pi f}) = 4 + 2\cos 2\pi f = 4 + e^{j2\pi f} + e^{-j2\pi f}$

$r_x(\tau) = \mathcal{F}^{-1}[P_x(e^{j2\pi f})] = 4\delta(\tau) + \delta(\tau+1) + \delta(\tau-1)$

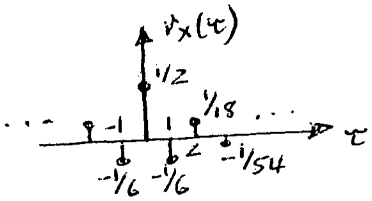


b) $P_x(e^{j2\pi f}) = \dfrac{2}{5 + 3\cos 2\pi f} \Rightarrow P_x(z) = \dfrac{2}{5 + \frac{3}{2}[z + z^{-1}]}$

$= \dfrac{4z}{(z+3)(3z+1)}$

$= \dfrac{\frac{3}{2}}{z+3} - \dfrac{\frac{1}{2}}{3z+1}$ ROC $\frac{1}{3} < |z| < 3$

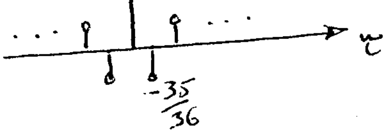
$r_x(\tau) = \frac{1}{2}(-3)^k u(-k) + \frac{1}{2}\left(-\frac{1}{3}\right)^k u(k-1) = \frac{1}{2}\left(-\frac{1}{3}\right)^{|k|}$



c) $P_x(z) = \dfrac{-4z^2}{(3z+1)(z+3)} + \dfrac{10z}{(3z+1)(z+3)} - \dfrac{4}{(3z+1)(z+3)}$ ROC $\frac{1}{3} < |z| < 3$

Recognizing from b) $\left(-\frac{1}{3}\right)^{|k|} \leftrightarrow \dfrac{8z}{(3z+1)(z+3)}$

$r_x(\tau) = -\frac{1}{2}\left(-\frac{1}{3}\right)^{|k+1|} + \frac{5}{4}\left(-\frac{1}{3}\right)^{|k|} - \frac{1}{2}\left(-\frac{1}{3}\right)^{|k-1|}$ (9)



Note maintenance of symmetry.

25/25

J. C.

2 (i) $E\{x[n]\} = \mu_x$

$$E\{x[n]x[m]\} = E\{x[n]x[n+\tau]\} = r_x(\tau)$$
$$\tau = |n-m|$$
$$c_x(0) = r_x(0) - \mu_x^2 < \infty.$$ (3)

(ii) a) $x[n] = \Theta \qquad \Theta \sim p_{\Phi}(\theta),\qquad E\{x[n]\} = E\{\Theta\}$ - Constant

$E\{x[n]x[m]\} = E\{\Theta^2\}$ - Constant

Assume $E\{(\Theta - E\{\Theta\})^2\} = c_x(0) < \infty$, then $x[n]$ is WSS.
But, $\lim_{N\to\infty} \frac{1}{N}\sum_{\tau=0}^{N-1} c_x(\tau) = c_x(0) \neq 0$, hence $x[n]$ is only ergodic in the mean if the variance of $\Theta = 0$, i.e. the pdf of Θ collapses to a delta function. (5)

b) $\mu_x = E\{x[n]\} = A E\{\cos(2\pi n f_0 + \phi)\}$

$$= A\int_{-\pi}^{\pi} \cos(2\pi n f_0 + \phi)\, d\phi = 0$$ - Constant

$$r_x(n,m) = E\{x[n]x[m]\} = A^2 E\{\cos(2\pi n f_0 + \phi)\cos(2\pi m f_0 + \phi)\}$$
$$= \frac{A^2}{2} E\{\cos(2\pi[n-m]f_0) + \cos(2\pi[n+m]f_0 + 2\phi)\}$$
$$= \frac{A^2}{2}\cos(2\pi\tau f_0) = r_x(\tau) = c_x(\tau)$$ - Function of τ only

$c_x(0) = \frac{A^2}{2} < \infty$, hence $x[n]$ is WSS

Checking for ergodicity in the mean,

$$\frac{1}{N}\sum_{\tau=0}^{N-1} c_x(\tau) = \frac{A^2}{2N}\sum_{\tau=0}^{N-1}\cos 2\pi\tau f_0 = \frac{A^2}{4N}\sum_{\tau=0}^{N-1}\left(e^{j2\pi f_0\tau} + e^{-j2\pi f_0 \tau}\right)$$
$$= \frac{A^2}{4N}\left\{\frac{1-e^{j2\pi N f_0}}{1-e^{j2\pi f_0}} + \frac{1-e^{-j2\pi N f_0}}{1-e^{-j2\pi f_0}}\right\}$$
$$= \frac{A^2}{2N}\frac{\sin(N\pi f_0)}{\sin(\pi f_0)}\cos([N-1]\pi f_0) = 0 \quad \lim_{N\to\infty}$$ provided $f_0 \neq 0$,

else $x[n] = A\cos(\phi)$, $c_x(\tau) = A^2/2$ and $x[n]$ is not ergodic in the mean. Therefore, $x[n]$ is ergodic in the mean provided $f_0 \neq 0$. (5)

c) $\mu_x = E\{x[n]\} = p - (1-p) = 2p - 1$ - Constant

$$r_x(n,m) = E\{x[n]x[m]\} = \begin{cases} E\{x^2[n]\} & n = m \\ E\{x[n]x[m]\} & n \neq m \end{cases}$$

AH J.C.

2) Cont.

(ii) c) $r_x(n,m) = \begin{cases} p + (1-p) = 1 & m = n \\ (1-2p)^2 & m \neq n \end{cases}$

$r_x(\tau) = 4p(1-p)\delta(\tau) + (1-2p)^2$
$r_x(0) = 1 < \infty$, hence $x[n]$ is WSS

Note $pdf(x[n], x[m]) = p^2\delta(x[n]-1, x[m]-1)$
$+ (1-p)^2\delta(x[n]+1, x[m]+1)$
$+ p(1-p)\delta(x[n]-1, x[m]+1)$
$+ p(1-p)\delta(x[n]+1, x[m]-1)$

As $c_x(\tau) = r_x(\tau) - \mu_x^2$
$= 4p(1-p)\delta(\tau)$

$\frac{1}{N}\sum_{\tau=0}^{N-1} c_x(\tau) = \frac{4p(1-p)}{N} \to 0$ as $N \to \infty$,

$x[n]$ is ergodic in the mean.

(12)

(25/25)

JC J.C.

3 (i) BLUE – Best Linear Unbiased Estimator

Restricts estimator to be linear in data, $x[n]$,

$$\hat{\theta}_i = \sum_{n=0}^{N-1} a_{in} x[n] \qquad i = 1, 2, \ldots, p$$

Parameters to be estimated

Best – minimum variance and unbiased will be equivalent to MVUE only when that turns out to be linear.

Only requires first two moments of the data. 5

(ii) $$E\{\hat{\theta}_i\} = \sum_{n=0}^{N-1} a_{in} E\{x[n]\} = \theta_i \qquad i = 1, 2, \ldots, p$$

In matrix form

$E\{\underline{\hat{\theta}}\} = A E\{\underline{x}\} = \underline{\theta}$ to be unbiased; from model of observation $E\{\underline{x}\} = H\underline{\theta}$

Thus

$$AH = I, \quad \text{with } A = \begin{bmatrix} \underline{a}_1^T \\ \underline{a}_2^T \\ \vdots \\ \underline{a}_p^T \end{bmatrix}, \; H = [\underline{h}_1 \underline{h}_2 \cdots \underline{h}_p],$$

this yields

$$\underline{a}_i^T \underline{h}_j = \delta_{ij} \qquad i = 1, 2, \ldots, p; \; j = 1, 2, \ldots, p \quad \text{– constraints}$$

$$\text{var}\{\hat{\theta}_i\} = E\{(\underline{a}_i^T(\underline{x} - E\{\underline{x}\}))^2\} = \underline{a}_i^T C \underline{a}_i$$

Form Lagrangian function, $J_i = \underline{a}_i^T C \underline{a}_i + \sum_{j=1}^{p} \lambda_j^{(i)} (\underline{a}_i^T \underline{h}_j - \delta_{ij})$

$$\frac{\partial J_i}{\partial \underline{a}_i} = 2 C \underline{a}_i + H \underline{\lambda}_i \quad \text{where } \underline{\lambda}_i = [\lambda_1^{(i)} \lambda_2^{(i)} \ldots \lambda_p^{(i)}]^T$$

Setting $\frac{\partial J_i}{\partial \underline{a}_i} = \underline{0} \implies \underline{a}_i = -\frac{1}{2} C^{-1} H \underline{\lambda}_i$

From the constraints, $H^T \underline{a}_i = \underline{e}_i$, where $\underline{e}_i = [0 \ldots 0 \underset{i\text{-th position}}{1} 0 \ldots 0]^T$,

thus $H^T \underline{a}_i = -\frac{1}{2} H^T C^{-1} H \underline{\lambda}_i = \underline{e}_i \implies -\frac{1}{2}\underline{\lambda}_i = (H^T C^{-1} H)^{-1} \underline{e}_i$,

and $\underline{a}_{i\,opt} = C^{-1} H (H^T C^{-1} H)^{-1} \underline{e}_i$.

Finally,

$$\underline{\hat{\theta}} = \begin{bmatrix} \underline{a}_{1opt}^T \underline{x} \\ \underline{a}_{2opt}^T \underline{x} \\ \vdots \\ \underline{a}_{popt}^T \underline{x} \end{bmatrix} = \begin{bmatrix} \underline{e}_1^T \\ \underline{e}_2^T \\ \vdots \\ \underline{e}_p^T \end{bmatrix} (H^T C^{-1} H)^{-1} H^T C^{-1} \underline{x} = (H^T C^{-1} H)^{-1} H^T C^{-1} \underline{x}$$

■

10

Ak J.C.

308 6/10

3) Cont.

(iii) $\underline{x} = H\underline{\theta} + \underline{w}$

B^{-1} exists, hence $\underline{\theta} = B^{-1}(\underline{\alpha} - \underline{b})$

$\Rightarrow \underline{x} = HB^{-1}(\underline{\alpha} - \underline{b}) + \underline{w}$, thus

$$\underbrace{\underline{x} + HB^{-1}\underline{b}}_{\underline{x}'} = \underbrace{HB^{-1}}_{H'}\underline{\alpha} + \underline{w},$$

from (ii) $\hat{\underline{\alpha}} = (H'^T C^{-1} H')^{-1} H'^T C^{-1} \underline{x}'$

$$= (B^{-1T} H^T C^{-1} H B^{-1})^{-1} B^{-1} H^T C^{-1} (\underline{x} + HB^{-1}\underline{b})$$

$$= B(H^T C^{-1} H)^{-1} H^T C^{-1} (\underline{x} + HB^{-1}\underline{b})$$

$$= B\hat{\underline{\theta}} + BB^{-1}\underline{b} = B\hat{\underline{\theta}} + \underline{b}$$

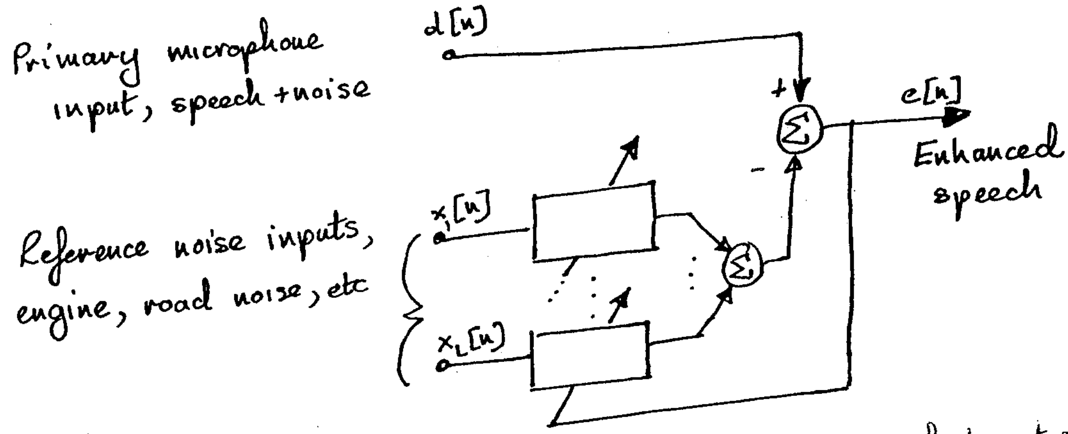
Hence $\hat{\underline{\alpha}} = B\hat{\underline{\theta}} + \underline{b}$ ■

10

25/25

AU J.C.

4) (i) Adaptive filter is used to enhance SNR at microphone input to speech recognition system. Extra noise reference inputs are provided by remote microphones.



(6)

(ii) $\underline{w}[n+1] = \underline{w}[n] + \mu\left(-\nabla_{\underline{w}} J\right)\Big|_{\underline{w}=\underline{w}[n]}$, $J = E\{e^2[n]\}$ method of steepest descent, drop $E\{\cdot\}$ in LMS to use instantaneous error squared,

$$\frac{\partial e^2[n]}{\partial \underline{w}} = 2e[n]\frac{\partial}{\partial \underline{w}}\left(d[n] - \underline{w}^T\underline{x}[n]\right) = -2e[n]\underline{x}[n]$$

$$\left.\begin{aligned} \underline{w}[n+1] &= \underline{w}[n] + 2\mu e[n]\underline{x}[n] \\ e[n] &= d[n] - \underline{w}^T\underline{x}[n] \end{aligned}\right\} \text{LMS Algorithm}$$

(5)

(iii) Need $J(\underline{w}_{\text{Wiener}})$ where $J = E\{e^2[n]\}$

$$J = \sigma_d^2 - 2\underline{p}^T\underline{w} + \underline{w}^T R\underline{w} \quad \text{where} \quad \left.\begin{aligned} R &= E\{\underline{x}\,\underline{x}^T\} \\ \underline{p} &= E\{\underline{x}\,d\} \end{aligned}\right\} \underline{x}, d \text{ jointly WSS.}$$

$\underline{w}_{\text{Wiener}}$ found from $\dfrac{\partial J}{\partial \underline{w}} = \underline{0}$

$$\Rightarrow -2\underline{p} + 2R\underline{w}_{\text{Wiener}} = \underline{0} \Rightarrow \underline{w}_{\text{Wiener}} = R^{-1}\underline{p}$$

Therefore
$$\begin{aligned} J_{MIN} &= \sigma_d^2 - 2\underline{p}^T R^{-1}\underline{p} + \underline{p}^T R^{-1} R R^{-1}\underline{p} \\ &= \sigma_d^2 - \underline{p}^T R^{-1}\underline{p} \end{aligned}$$

(7)

Gradient noise in LMS will introduce excess MSE, $J_{ex}(\infty)$, hence non zero misadjustment $\mathcal{M} \triangleq \dfrac{J_{ex}(\infty)}{J_{MIN}}$

J.C.

4) (iv)

a) $\nabla_{\underline{w}} J \big|_{\underline{w} = \underline{w}[n]}$

$$\frac{\partial\left[\lambda e^2[n] + (1-\lambda)|e[n]|\right]}{\partial \underline{w}} = -2e[n]\underline{x}[n]\lambda - \text{sign}(e[n])\underline{x}[n](1-\lambda)$$

$$\underline{w}[n+1] = \underline{w}[n] + \mu\left(2e[n]\lambda + \text{sign}(e[n])(1-\lambda)\right)\underline{x}[n]$$

b) Advantages: Robustness to impulsive noise in desired response
Combines LMS / Least Absolute Error algorithms

Disadvantages: Slower convergence than LMS except when $\lambda = 1.0$.
Higher computational complexity.

c) Assume desired response has Gaussian distribution, estimate variance $\hat{\sigma}_d^2 = \sum_{k=n-L+1}^{n} d^2[k]$ over a sliding window, if instantaneous $d^2[k] >> \hat{\sigma}_d^2$ then $\lambda \to 0$, i.e. use LAE algo., else $\lambda \to 1$, use LMS algorithm.

(7)

25/25

JC.

5) (i) Block-based – estimator needs entire observation vector to be collected before it can be calculated

e.g. sample mean $\hat{\mu} = \frac{1}{N}\sum_{n=0}^{N-1} x[n]$

Sequential – estimate is refined as each new sample arises

e.g. $\hat{\mu}[N] = \hat{\mu}[N-1] + \frac{1}{N+1}\left[x[N] - \hat{\mu}[N-1]\right]$ (3)

(ii) $\underline{\epsilon} = (\underline{x} - \underline{s})$ is $\perp$ to the columns of H, write

$H = [\underline{h}_1\ \underline{h}_2 \cdots \underline{h}_p]$, $\underline{\epsilon}^T \underline{h}_i = 0$ for $i = 1, 2, \ldots, p$, when $\hat{\underline{\theta}} = \hat{\underline{\theta}}_{LS}$ (4)

(iii) $\underline{\epsilon}^T[\underline{h}_1\ \underline{h}_2 \cdots \underline{h}_p] = \underline{0}^T$

$\Rightarrow (\underline{x} - H\hat{\underline{\theta}}_{LS})^T H = \underline{0}^T$

$\Rightarrow \underline{x}^T H - \hat{\underline{\theta}}_{LS}^T H^T H = \underline{0}^T$

$\Rightarrow H^T\underline{x} - H^T H\hat{\underline{\theta}}_{LS} = \underline{0}$ $\quad\Rightarrow \hat{\underline{\theta}}_{LS} = (H^T H)^{-1} H^T \underline{x}$. ■ (5)

(iv) $J_{MIN} = (\underline{x} - H\hat{\underline{\theta}}_{LS})^T(\underline{x} - H\hat{\underline{\theta}}_{LS})$.

$= (\underline{x} - H\hat{\underline{\theta}}_{LS})^T \underline{x}$ from $\perp$ condition.

$= (\underline{x} - H(H^T H)^{-1} H^T \underline{x})^T \underline{x}$ (3)

$= \underline{x}^T (I - H(H^T H)^{-1} H^T)\underline{x}$ ■

(v) $\hat{\underline{\theta}}[n] = (H^T[n] H[n])^{-1} H^T[n]\underline{x}[n] = \left([H^T[n-1]\ \underline{h}[n]] \begin{bmatrix} H[n-1] \\ \underline{h}^T[n] \end{bmatrix} \right)^{-1} \left([H^T[n-1]\ \underline{h}[n]] \begin{bmatrix} \underline{x}[n-1] \\ x[n] \end{bmatrix} \right)$

$= \left(H^T[n-1]H[n-1] + \underline{h}[n]\underline{h}^T[n]\right)^{-1}\left(H^T[n-1]\underline{x}[n-1] + \underline{h}[n]x[n]\right)$

Let $\Sigma[n-1] = (H^T[n-1]H[n-1])^{-1}$ – covariance matrix of $\hat{\underline{\theta}}[n-1]$

$\hat{\underline{\theta}}[n] = \left(\Sigma^{-1}[n-1] + \underline{h}[n]\underline{h}^T[n]\right)^{-1}\left(H^T[n-1]\underline{x}[n-1] + \underline{h}[n]x[n]\right)$

$\Sigma[n] = \left(\Sigma^{-1}[n-1] + \underline{h}[n]\underline{h}^T[n]\right)^{-1} = \Sigma[n-1] - \dfrac{\Sigma[n-1]\underline{h}[n]\underline{h}^T[n]\Sigma[n-1]}{1 + \underline{h}^T[n]\Sigma[n-1]\underline{h}[n]}$

JC

i) (v) Cont.

$$\underline{\Sigma}[n] = (\underline{I} - \underline{K}[n]\underline{h}^T[n])\underline{\Sigma}[n-1]$$

where the Kalman gain vector

$$\underline{K}[n] = \frac{\underline{\Sigma}[n-1]\underline{h}[n]}{1+\underline{h}^T[n]\underline{\Sigma}[n-1]\underline{h}[n]}$$

$$\hat{\underline{\theta}}[n] = (\underline{I}-\underline{K}[n]\underline{h}^T[n])\underline{\Sigma}[n-1](\underline{\Sigma}^{-1}[n-1]\hat{\underline{\theta}}[n-1]+\underline{h}[n]x[n])$$
$$= \hat{\underline{\theta}}[n-1] + \underline{\Sigma}[n-1]\underline{h}[n]x[n] - \underline{K}[n]\underline{h}^T[n]\hat{\underline{\theta}}[n-1] - \underline{K}[n]\underline{h}^T[n]\underline{\Sigma}[n-1]\underline{h}[n]x[n]$$

But

$$\underline{\Sigma}[n-1]\underline{h}[n] - \underline{K}[n]\underline{h}^T[n]\hat{\underline{\theta}}[n-1] = (1+\underline{h}^T[n]\underline{\Sigma}[n-1]\underline{h}[n])\underline{k}[n] - \underline{k}[n]\underline{h}^T[n]\underline{\Sigma}[n-1]\underline{h}[n] = \underline{k}[n]$$

Therefore $\hat{\underline{\theta}}_{LS}[n] = \hat{\underline{\theta}}_{LS}[n-1] + \underline{k}[n](x[n]-\underline{h}^T[n]\hat{\underline{\theta}}_{LS}[n-1])$ (10)

25/25

Ahh J.C.