2005 IIB 4F7 Dr S Godsill sjg@eng.cam.ac.uk

ENGINEERING TRIPOS PART IIB

Mon 2 May 2005 9 to 10.30

Module 4F7

DIGITAL FILTERS AND SPECTRUM ESTIMATION - WORKED SOLUTIONS

Answer not more than ***three*** *questions.*

All questions carry the same number of marks.

The ***approximate*** *percentage of marks allocated to each part of a question is indicated in the right margin.*

There are no attachments.

You may not start to read the questions printed on the subsequent pages of this question paper until instructed that you may do so by the Invigilator

(TURN OVER

1 Let $\{x(n)\}$ be a sequence of independent and identically distributed symbols such that

$$\Pr\{x(n)=1\}=\Pr\{x(n)=-1\}=0.5.$$

These symbols are transmitted through a communication channel and one observes

$$y(n)=\sum_{k=0}^{L-1}\alpha_k x(n-k)+v(n)$$

where $\{v(n)\}$ is zero-mean white noise of variance $E\left(v^2(n)\right)=\sigma_v^2$ and is statistically independent of $\{x(n)\}$.

(a) Determine the Wiener filter $\mathbf{h}_{\text{opt}}$ which minimizes

$$J(\mathbf{h})=E\left\{\left(x(n)-\mathbf{h}^{\mathrm{T}}\mathbf{y}(n)\right)^2\right\} \tag{1}$$

where

$$\mathbf{y}(n)=\left(y(n)\;\; y(n+1)\;\; \cdots\;\; y(n+M-1)\right)^{\mathrm{T}}$$

as a function of

$$\mathbf{R}=E\left[\mathbf{y}(n)\mathbf{y}^{\mathrm{T}}(n)\right],$$
$$\mathbf{p}=E\left[\mathbf{y}(n)x(n)\right].$$

Compute $\mathbf{R}$ and $\mathbf{p}$ explicitly. [50%]

Answer. We have

$$\begin{aligned}J(\mathbf{h})&=E\left\{\left(x(n)-\mathbf{h}^{\mathrm{T}}\mathbf{y}(n)\right)^2\right\}\\&=E\left\{x^2(n)\right\}-2\mathbf{h}^{\mathrm{T}}E\left\{(x(n)\mathbf{y}(n))\right\}+\mathbf{h}^{\mathrm{T}}E\left\{\mathbf{y}(n)\mathbf{y}(n)^{\mathrm{T}}\right\}\mathbf{h}.\end{aligned}$$

By taking the derivative with respect to $\mathbf{h}$, we obtain

$$\begin{aligned}\mathbf{h}_{\text{opt}}(M)&=E\left\{\mathbf{y}(n)\mathbf{y}(n)^{\mathrm{T}}\right\}^{-1}E\left\{(x(n)\mathbf{y}(n))\right\}\\&=\mathbf{R}^{-1}\mathbf{p}\end{aligned}$$

We also have

$$E\left[y(n)x(n)\right]=\sum_{k=0}^{L-1}\alpha_k E\left[x(n-k)x(n)\right]+E\left[v(n)x(n)\right]=\alpha_0$$

(cont.

as

$$E\left[x\left(n-k\right)x\left(n\right)\right]=\delta\left(k\right).$$

Similarly

$$E\left[y\left(n+1\right)x\left(n\right)\right]=\sum_{k=0}^{L-1}\alpha_{k}E\left[x\left(n+1-k\right)x\left(n\right)\right]+E\left[v\left(n+1\right)x\left(n\right)\right]=\alpha_{1}$$

and one can show

$$E\left[y\left(n+k\right)x\left(n\right)\right]=\begin{cases}\alpha_{k} & \text{if } k\leq L-1,\\ 0 & \text{otherwise.}\end{cases}$$

We also have

$$\begin{aligned}
&E\left(y\left(n+m\right)y\left(n+l\right)\right)\\
&=\sum_{k=0}^{L-1}\sum_{j=0}^{L-1}E\left(\left(\alpha_{k}x\left(n+m-k\right)+v\left(n+m\right)\right)\left(\alpha_{j}x\left(n+l-j\right)+v\left(n+l\right)\right)\right)\\
&=\sum_{k=0}^{L-1}\sum_{j=0}^{L-1}\alpha_{k}\alpha_{j}E\left(x\left(n+m-k\right)x\left(n+l-j\right)\right)+\sigma_{v}^{2}\delta\left(m-l\right)\\
&=\sum_{k=0}^{L-1}\sum_{j=0}^{L-1}\alpha_{k}\alpha_{j}\delta\left(m-k-l+j\right)+\sigma_{v}^{2}\delta\left(m-l\right).
\end{aligned}$$

as

$$E\left(x\left(n+m-k\right)x\left(n+l-j\right)\right)=\delta\left(m-k-l+j\right)$$

(b) The value $J\left(\mathbf{h}_{\text{opt}}\right)$ is actually a function of the length M of the Wiener filter. Without any calculation, explain why there exists a value K such that this function is constant for $M\geq K$. What is the value of K? [25%]

Answer. As a function of M, $J\left(\mathbf{h}_{\text{opt}}\right)$ is constant and reaches its minimum for $M\geq L$. This simply follows from the fact that the observations

$$\left(y\left(n+L\right),y\left(n+L+1\right),y\left(n+L+2\right)\cdots\right)$$

are statistically independent of $x\left(n\right)$. So if one increases the channel length, no further information about $x\left(n\right)$ can be gathered.

(c) In practical situations, the Wiener filter cannot be computed as it relies on the coefficients $\{\alpha_k\}$ and σ_v^2 which are unknown quantities. Using a training sequence of symbols $\{x(n)\}$, describe how you would apply the LMS algorithm in this context to approximate the Wiener filter. What are the advantages and disadvantages of such an approach? [25%]

Answer. In practice, one uses a training sequence known to the receiver. Then the LMS algorithm is used to approximate the Wiener filter with $y(n)$ as an input and $x(n)$ as the reference signal. The advantages of such an approach are that the calculations can be performed online, $\{\alpha_k\}$ and σ_v^2 do not need to be known and the algorithm can adapt itself to a non-stationary environment. The main drawback of this approach is that it is bandwidth consuming as it is necessary to transmit a training sequence.

2 (a) Consider the following signals

$$\begin{array}{ll} L\text{-tap FIR Filter} & u(n) = \sum_{k=0}^{L-1} \beta_k w(n-k) \\ 2\text{-tap IIR Filter} & u(n) = \alpha u(n-1) + w(n) \end{array}$$

where $|\alpha| < 1$ and $\{w(n)\}$ is zero-mean white noise of variance $E\left(w^2(n)\right) = \sigma^2$.

If these signals are the input to an LMS filter of length L, what is the stability limit on the stepsize μ given by $\left(ME\left\{u^2(n)\right\}\right)^{-1}$ for these two signals?

[30%]

Answer. In both cases, we need to compute $E\left\{u^2(n)\right\}$. For the FIR filter, we have

$$\begin{aligned} E\left\{u^2(n)\right\} &= E\left\{\left(\sum_{k=0}^{L-1} \beta_k w(n-k)\right)^2\right\} \\ &= \sum_{k=0}^{L-1} \beta_k^2 E\left\{w^2(n-k)\right\} \\ &= \left(\sum_{k=0}^{L-1} \beta_k^2\right) \sigma^2. \end{aligned}$$

For the IIR filter, we have

$$\begin{aligned} u^2(n) &= (\alpha u(n-1) + w(n))^2 \\ &= \alpha^2 u^2(n-1) + w^2(n) + 2\alpha u(n-1) w(n). \end{aligned}$$

Thus it follows that

$$E\left(u^2(n)\right) = \frac{\sigma^2}{1-\alpha^2}.$$

(b) Let $\{u(n)\}$ be a zero-mean input signal and $\{d(n)\}$ be a reference signal. Consider the following recursive algorithm

$$\mathbf{h}(n) = (1-\mu\gamma)\,\mathbf{h}(n-1) + \mu \mathbf{u}(n)\, e(n) \qquad (2)$$

where

$$e(n) = d(n) - \mathbf{h}^{\mathrm{T}}(n-1)\,\mathbf{u}(n)$$

(CONTINUED OVER.

with $\gamma > 0$, $\mathbf{u}(n) = (u(n)\ u(n-1) \cdots\ u(n-M+1))^{\mathrm{T}}$ and $\mathbf{h}(n) = (h(n)\ h(n-1) \cdots\ h(n-M+1))^{\mathrm{T}}$.

Assuming the expectation of $\mathbf{h}(n)$, denoted $E[\mathbf{h}(n)]$, converges towards a limit $\mathbf{h}$, determine this limit by making the following standard approximation

$$E\left[\mathbf{u}(n)\mathbf{u}^{\mathrm{T}}(n)\mathbf{h}(n-1)\right] \approx E\left[\mathbf{u}(n)\mathbf{u}^{\mathrm{T}}(n)\right] E[\mathbf{h}(n-1)]. \tag{3}$$

Express the result as a function of

$$\mathbf{R} = E\left[\mathbf{u}(n)\mathbf{u}^{\mathrm{T}}(n)\right],$$
$$\mathbf{p} = E[\mathbf{u}(n)d(n)].$$

[40%]

Answer. Clearly one has

$$\begin{aligned} E[\mathbf{h}(n)] &= (1-\mu\gamma)E[\mathbf{h}(n-1)] + \mu E[\mathbf{u}(n)e(n)] \\ &= (1-\mu\gamma)E[\mathbf{h}(n-1)] + \mu E[\mathbf{u}(n)d(n)] - \mu E\left[\mathbf{u}(n)\mathbf{u}^{\mathrm{T}}(n)\mathbf{h}(n-1)\right] \\ &\approx (1-\mu\gamma)E[\mathbf{h}(n-1)] + \mu E[\mathbf{u}(n)d(n)] - \mu E\left[\mathbf{u}(n)\mathbf{u}^{\mathrm{T}}(n)\right] E[\mathbf{h}(n-1)]. \end{aligned}$$

If there is a limit, it satisfies

$$\mathbf{h} = (1-\mu\gamma)\mathbf{h} + \mu\mathbf{p} - \mu\mathbf{R}\mathbf{h}.$$

Thus it follows that

$$\mathbf{h} = (\gamma\mathbf{I} + \mathbf{R})^{-1}\mathbf{p}(n).$$

(c) An approximate analysis of the recursion (2) shows that it is numerically stable if

$$\mu \le \frac{2}{\lambda_{\max} + \gamma}$$

where $\lambda_{\max}$ is the largest eigenvalue of $\mathbf{R}$. In practice this eigenvalue cannot be computed exactly. Propose and justify an alternative criterion based on $E\left\{u^2(n)\right\}$ ensuring stability of (2). How would you approximate $E\left\{u^2(n)\right\}$ in a real-world scenario? [30%]

(cont.

Answer. We have

$$\begin{aligned}\text{trace } \mathbf{R} &= \sum_{i=1}^{M} \lambda_i \\ &= ME\left\{u^2(n)\right\} \\ &\le M\lambda_{\max}.\end{aligned}$$

So we can use

$$\mu \le \frac{2}{ME\left\{u^2(n)\right\} + \gamma}.$$

$E\left\{u^2(n)\right\}$ can be approximated numerically by averaging the signal $\{u(n)\}$ over a moving window.

3 (a) The periodogram estimate for power spectrum estimation is defined as

$$\hat{S}_X(e^{j\omega T}) = \sum_{k=-(N-1)}^{N-1} \hat{R}_{XX}[k]\, e^{-jk\omega T}$$

Discuss briefly the properties of the periodogram, including bias, variance, frequency resolution and computation.

Show that when the *biased* estimator for autocorrelation is applied, the periodogram can be determined directly from the DTFT of a windowed version of the original data,

$$\hat{S}_X(e^{j\omega T}) = \frac{1}{N}|X_w(e^{j\omega T})|^2$$

where $X_w(e^{j\omega T})$ should be carefully defined. [50%]

Answer: Periodogram is biased, but asymptotically unbiased. The variance can be very high for random signals and doesn't necessarily improve with data length. Frequency resolution determined by data length, computation usually efficiently carried out using FFT.

To see this, rewrite the biased estimate as:

$$\begin{aligned}\hat{R}_{XX}[k] &= \frac{1}{N}\sum_{n=0}^{N-1-k} x_n\, x_{n+k} \\ &= \frac{1}{N}\sum_{n=-\infty}^{\infty} v_n\, v_{n+k}\end{aligned}$$

where $v_n = w_n x_n$ is a version of x_n truncated by multiplication with a rectangular window:

$$w_n = \begin{cases} 1, & n = 0, 1, ..., N-1 \\ 0, & \text{otherwise} \end{cases}$$

Now, letting $n' = -n$ and $u_n = v_{-n}$ (time-reversal), we have:

$$\begin{aligned} \hat{R}_{XX}[k] &= \frac{1}{N} \sum_{n'=-\infty}^{\infty} v_{-n'} v_{(-n'+k)} \\ &= \frac{1}{N} \sum_{n'=-\infty}^{\infty} u_{n'} v_{k-n'} = \frac{1}{N} u_k * v_k \end{aligned}$$

i.e. a standard discrete time convolution of u_k with v_k. Taking the DTFT of both sides we get (by the discrete time convolution theorem):

$$\hat{S}_X(e^{j\omega T}) = \frac{1}{N} U(e^{j\omega T}) V(e^{j\omega T})$$

where:

$$\begin{aligned} V(e^{j\omega T}) &= \sum_{n=-\infty}^{+\infty} v_n e^{-jn\omega T} = \sum_{n=0}^{(N-1)} x_n e^{-jn\omega T} \\ &= X_w(e^{j\omega T}) \end{aligned}$$

and

$$\begin{aligned} U(e^{j\omega T}) &= \sum_{n=-\infty}^{+\infty} u_n e^{-jn\omega T} = \sum_{n=0}^{1-N} x_{-n} e^{-jn\omega T} \\ &= \sum_{n=0}^{N-1} x_n e^{+jn\omega T} = X_w^*(e^{j\omega T}) \end{aligned}$$

where X_w is the DTFT of the windowed signal $x_n w_n$.

Hence

$$\begin{aligned} \hat{S}_X(e^{j\omega T}) &= \frac{1}{N} U(e^{j\omega T}) V(e^{j\omega T}) \\ &= \frac{1}{N} X_w^*(e^{j\omega T}) X_w(e^{j\omega T}) = \frac{1}{N} |X_w(e^{j\omega T})|^2 \end{aligned}$$

(cont.

(b) A moving average random process is defined as:

$$x_n = 0.9e_n + 0.1e_{n-1}$$

where e_n is white noise having unity variance. Determine the power spectrum of this process and sketch it over the normalised frequency range 0 to 2π .

The biased method of autocorrelation estimation is employed. Show that the expected value of the corresponding periodogram estimate of the power spectrum is given by

$$\hat{S}_X(e^{j\omega T}) = 0.82 + 0.18\frac{N-1}{N}\cos(\omega T)$$

Is the periodogram estimate biased for this MA process? Is it biased asymptotically (i.e. as the data size becomes very large)? You should carefully explain your answer

[50%]

Answer:

Autocorrelation function:

$$R_{XX}[l] = E[x_n x_{n+l}] = 0.82\delta_l + 0.09\delta_{l-1} + 0.09\delta_{l+1}$$

since $E[e_n e_{n+m}] = \delta_m$.

Therefore power spectrum is given by:

$$S_X(e^{j\omega T}) = 0.82 + 0.18\cos(\omega T)$$

Consider the periodogram:

$$\begin{aligned}
\hat{S}_X(e^{j\omega T}) &= \frac{1}{N}|X_w(e^{j\omega T})|^2 \\
&= \frac{1}{N}|\sum_{n=0}^{N-1} x_n e^{-j\omega nT})|^2 \\
&= \frac{1}{N}(\sum_{n_1=0}^{N-1} x_{n_1} e^{-j\omega n_1 T})(\sum_{n_2=0}^{N-1} x_{n_2} e^{+j\omega n_2 T}) \\
&= \frac{1}{N}(\sum_{n_1=0}^{N-1}\sum_{n_2=0}^{N-1} x_{n_1} x_{n_2} e^{-j\omega(n_1-n_2)T})
\end{aligned}$$

(CONTINUED OVER.

and taking expectations

$$E[\hat{S}_X(e^{j\omega T})] = \frac{1}{N}(\sum_{n_1=0}^{N-1}\sum_{n_2=0}^{N-1} R_{XX}(n_2 - n_1)e^{-j\omega(n_1-n_2)T})$$
$$= 0.82 + 0.09\frac{N-1}{N}e^{-j\omega T} + 0.09\frac{N-1}{N}e^{+j\omega T}$$
$$= 0.82 + 0.18\frac{N-1}{N}cos(\omega T)$$

Hence biased, since $E[\hat{S}_X(e^{j\omega T})] \neq S_X(e^{j\omega T})$. However, asymptotically unbiased since $(N-1)/N \to 1$ as $N \to \infty$.

4 (a) An ARMA model has the following difference equation:

$$x_n = -\sum_{p=1}^{P} a_p \, x_{n-p} + \sum_{q=0}^{Q} b_q \, w_{n-q}$$

Discuss how the ARMA model can be used for power spectrum estimation. You should include the formula for the power spectrum of the ARMA model, as well as a discussion of its advantages/disadvantages compared with non-parametric procedures. [35%]

Answer: Bookwork, taken from:

ARMA Models A quite general representation is the autoregressive moving-average (ARMA) model:

- The ARMA(P,Q) model difference equation representation is:

$$\boxed{x_n = -\sum_{p=1}^{P} a_p \, x_{n-p} + \sum_{q=0}^{Q} b_q \, w_{n-q}} \tag{4}$$

where:

a_p are the AR parameters,

b_q are the MA parameters

and $\{W_n\}$ is a zero-mean stationary white noise process with unit variance, $\sigma_w^2 = 1$.

- Clearly the ARMA model is a pole-zero IIR filter-based model with transfer function

$$H(z) = \frac{B(z)}{A(z)}$$

where:

$$A(z) = 1 + \sum_{p=1}^{P} a_p z^{-p}, \quad B(z) = \sum_{q=0}^{Q} b_q z^{-q}$$

- Unless otherwise stated we will always assume that the filter is stable, i.e. the poles (solutions of $A(z) = 0$) all lie *within* the unit circle (we say in this case that $A(z)$ is *minimum phase*). Otherwise the autocorrelation function is undefined and the process is technically *non-stationary*.

•Hence the power spectrum of the ARMA process is:

$$S_X(e^{j\omega T}) = \frac{|B(e^{j\omega T})|^2}{|A(e^{j\omega T})|^2}$$

Thus, estimate the parameters a and b from the data, then plug into spectral density formula.

The ARMA model is quite a flexible and general way to model a stationary random process:

•The poles model well the *peaks* in the spectrum (sharper peaks implies poles closer to the unit circle)

•The zeros model troughs in the spectrum

•Complex spectra can be approximated well by large model orders P and Q

Can give improved variance of estimation; however, may be highly biased and inaccurate when an ARMA model is inappropriate for the data. Also, quite expensive to compute parameters accurately.

(b) Show that the ARMA model autocorrelation function obeys the following difference equation

$$R_{XX}[r] + \sum_{p=1}^{P} a_p\, R_{XX}[r-p] = \begin{cases} c_r & \text{if } r \leq Q \\ 0, & \text{if } r > Q \end{cases}$$

where:

$$c_r = \sum_{q=r}^{Q} b_q\, h_{q-r}$$

and h_n is the impulse response of the corresponding IIR filter. [35%]

Answer:

Autocorrelation function for ARMA Model The autocorrelation function $R_{XX}[r]$ for the output x_n of the ARMA model is:

$$R_{XX}[r] = E[x_n x_{n+r}]$$

(cont.

Substituting for x_{n+r} from equation 4 gives:

$$R_{XX}[r] = E\left[x_n\left\{-\sum_{p=1}^{P} a_p\, x_{n+r-p} + \sum_{q=0}^{Q} b_q\, w_{n+r-q}\right\}\right]$$

$$= -\sum_{p=1}^{P} a_p\, E[x_n x_{n+r-p}] + \sum_{q=0}^{Q} b_q\, E[x_n w_{n+r-q}]$$

The white noise process $\{W_n\}$ is wide-sense stationary so that $\{X_n\}$ is also wide-sense stationary provided the the ARMA filter is stable. Therefore:

$$\boxed{R_{XX}[r] = -\sum_{p=1}^{P} a_p\, R_{XX}[r-p] + \sum_{q=0}^{Q} b_q\, R_{XW}[r-q]} \qquad (5)$$

Note that the auto-correlation and cross-correlation satisfy the same ARMA system difference equation as x_n and w_n .

The cross-correlation term $R_{XW}[.]$ can be obtained as follows. Let the system impulse response be h_n , then:

$$x_n = \sum_{m=\infty}^{\infty} h_m\, w_{n-m}$$

Therefore,

$$E[x_n\, w_{n+k}] = E[w_{n+k} \sum_{m=\infty}^{\infty} h_m\, w_{n-m}]$$

$$R_{XW}[k] = \sum_{m=-\infty}^{\infty} h_m\, E[w_{n+k}\, w_{n-m}]$$

Now the noise is a zero-mean stationary white process so that:

$$E[w_{n+k}\, w_{n-m}] = \begin{cases} \sigma_W^2 & \text{if } m = -k \\ 0 & \text{otherwise} \end{cases}$$

(CONTINUED OVER.

and $\sigma_W^2 = 1$ without loss of generality. Hence,

$$R_{XW}[k] = h_{-k}$$

Substituting this expression for $R_{XW}[k]$ into equation 5 gives the *Yule-Walker Equation* for an ARMA process,

$$\boxed{R_{XX}[r] = -\sum_{p=1}^{P} a_p \, R_{XX}[r-p] + \sum_{q=0}^{Q} b_q \, h_{q-r}} \qquad (6)$$

Since the system is causal, equation 6 may be rewritten as:

$$\boxed{R_{XX}[r] = -\sum_{p=1}^{P} a_p \, R_{XX}[r-p] + c_r} \qquad (7)$$

where:

$$c_r = \begin{cases} \sum_{q=r}^{Q} b_q \, h_{q-r} & \text{if } r \leq Q \\ 0 & \text{if } r > Q \end{cases} \qquad (8)$$

(c) Four values from the autocorrelation function of an ARMA model with $P = 2$ and $Q = 1$ are given by

$$R_{XX}[0] = 1.84, \;\; R_{XX}[1] = 1.32, \;\; R_{XX}[2] = 0.75, \;\; R_{XX}[3] = 0.47$$

Use the result of part b) to set up and solve the equations for the AR coefficients a_1 and a_2 in this ARMA model. [30%]

Answer: We can use the above result for $r = 2, 3$ to formulate a pair of simultaneous equations involving only the AR parameters:

$$R_{XX}[r] + \sum_{p=1}^{P} a_p \, R_{XX}[r-p] = 0, \;\; r = 2, 3$$

Solving these we get:

$$a_1 = 0.33, \;\; a_2 = 0.17$$

END OF PAPER