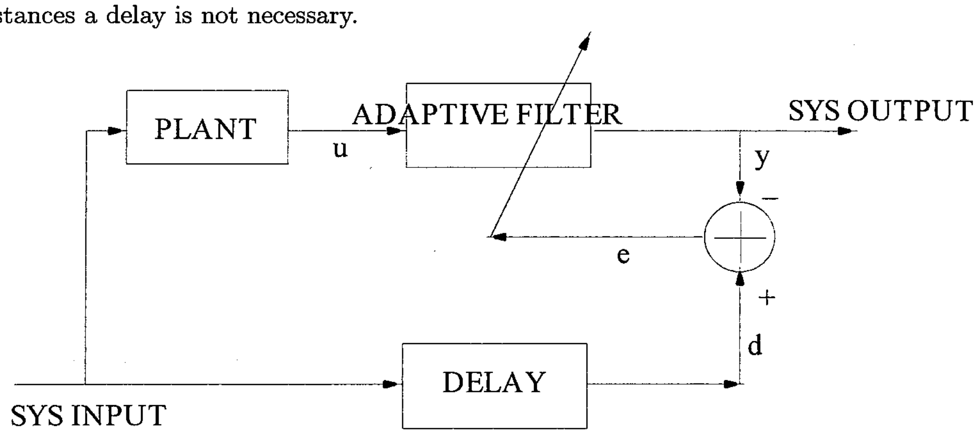
PARTIIB 2011 DR SS SINGH

Module 4F7 - Digital Filters and Spectrum Estimation
Friday 06 May 2011, 14:30-16:00

Question 1

Answers:

a) The desired response is a delayed version of the plant's input. In some instances a delay is not necessary.



b)

$$\mathbf{h}^* = \mathbf{h}^* + \mu\mathbf{R}^{-1}(\mathbf{p} - \mathbf{R}\mathbf{h}^*)$$
$$\mathbf{h}^* = \mathbf{R}^{-1}\mathbf{p}$$

c) Let $\mathbf{v}(n) = \mathbf{h}(n) - \mathbf{h}^*$. Thus

$$\begin{aligned}\mathbf{h}(n) - \mathbf{h}^* &= \mathbf{h}(n-1) - \mathbf{h}^* + \mu\mathbf{R}^{-1}(\mathbf{R}\mathbf{h}^* - \mathbf{R}\mathbf{h}(n-1))\\ \mathbf{v}(n) &= \mathbf{v}(n-1) - \mu\mathbf{v}(n-1)\\ &= (1-\mu)\mathbf{v}(n-1)\end{aligned}$$

$|1-\mu| < 1$ guarantees convergence. Optimal $\mu = 1$ as it would result in convergence in one step.

d) The modification of the LMS is

$$\begin{aligned}\mathbf{h}(n) &= \mathbf{h}(n-1) + \mu\mathbf{R}(n)^{-1}\mathbf{u}(n)(d(n) - \mathbf{u}(n)^{\mathrm{T}}\mathbf{h}(n-1))\\ \mathbf{R}(n) &= \frac{n}{n+1}\mathbf{R}(n-1) + \frac{1}{n+1}\mathbf{u}(n)\mathbf{u}(n)^{\mathrm{T}}\end{aligned}$$

and $\mathbf{R}(0) = e\mathbf{I}$, e a very small positive constant, to ensure invertibility.

From the study of $\mathbf{v}(n) = \mathbf{h}(n) - \mathbf{h}^*$ above we see that this algorithm will be insensitive to the eigenvalue spread of $\mathbf{R}$ whereas the LMS is sensitive.

e) For the slowly time varying statistics, use

$$\mathbf{R}(n) = \lambda\mathbf{R}(n-1) + \mathbf{u}(n)\mathbf{u}(n)^{\mathrm{T}}$$

where $0 < \lambda < 1$. Thus we have the recursive least squares algorithm

Comments: This was a very popular question. The application to inverse modeling was answered well by most although some gave the block diagram for identification instead. Part (b) on the limit point as well easily solved. For part (c), some failed to see that convergence is in one iteration when the step-size equals to 1. Part (d) proved challenging for many – in the modified LMS a better answer would have been to use the sample average of the autocorrelation function and not the instantaneous estimate. Many failed to comment that this new algorithm is less sensitive to the eigenvalue spread compared to the LMS. For part (e), many failed to see that an exponentially weighted estimate of the autocorrelation function would have been better than an instantaneous estimate.

Question 2

Answers:

a)

$$u(n) = a\,u(n-1) + v(n) + b\,v(n-1)$$

Multiply both sides of this equation by $u(n)$ and take the expectation. Do the same with $u(n-1)$ and $u(n-2)$.

$$\begin{aligned} r_0 &= ar_1 + E\left\{[v(n) + b\,v(n-1)]\,u(n)\right\} \\ r_1 &= ar_0 + E\left\{[v(n) + b\,v(n-1)]\,u(n-1)\right\} \\ r_2 &= ar_1 + E\left\{[v(n) + b\,v(n-1)]\,u(n-2)\right\} \end{aligned}$$

where

$$\begin{aligned} E\left\{[v(n) + b\,v(n-1)]\,u(n-2)\right\} &= 0 \\ E\left\{[v(n) + b\,v(n-1)]\,u(n-1)\right\} &= b\sigma_v^2 \end{aligned}$$

and

$$\begin{aligned} &E\left\{[v(n) + b\,v(n-1)]\,u(n)\right\} \\ &= E\left\{[v(n) + b\,v(n-1)]\,[a\,u(n-1) + v(n) + b\,v(n-1)]\right\} \\ &= ab\sigma_v^2 + \sigma_v^2(1+b^2) \end{aligned}$$

b)

$$E\left\{\left(d(n) - \mathbf{h}^{\mathrm{T}}\mathbf{u}(n)\right)\right\} = \sigma_d^2 + \mathbf{h}^{\mathrm{T}}\mathbf{R}\mathbf{h} - 2\mathbf{h}^{\mathrm{T}}\mathbf{p}.$$

Differentiating with respect to $\mathbf{h}$, the minimiser is $\mathbf{h} = \mathbf{R}^{-1}\mathbf{p}$ or $\mathbf{R}\mathbf{h} = \mathbf{p}$. Thus

$$\begin{aligned} J_{\min} - \sigma_d^2 &= \mathbf{h}^{\mathrm{T}}\mathbf{R}\mathbf{h} - 2\mathbf{h}^{\mathrm{T}}\mathbf{p} \\ &= -\mathbf{h}^{\mathrm{T}}\mathbf{p} \\ &= -\mathbf{p}^{\mathrm{T}}\left(\mathbf{R}^{-1}\right)^{\mathrm{T}}\mathbf{p} \\ &= -\mathbf{p}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{p} \end{aligned}$$

since $\left(\mathbf{R}^{-1}\right)^{\mathrm{T}} = \left(\mathbf{R}^{\mathrm{T}}\right)^{-1} = \mathbf{R}^{-1}$.

c) Since we are predicting, $d(n) = u(n)$.

For first order predictor, $\mathbf{u}(n) = u(n-1)$, $\mathbf{R} = r_0$, $\mathbf{p} = E\{u(n-1)d(n)\} = r_1$. Thus

$$J_{\min} = r_0 - \frac{r_1^2}{r_0}.$$

For an order two predictor, $\mathbf{u}(n) = [u(n-1), u(n-2)]^{\mathrm{T}}$,

$$\mathbf{R} = \begin{bmatrix} r_0 & r_1 \\ r_1 & r_0 \end{bmatrix}, \qquad \mathbf{p} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} r_1 \\ ar_1 \end{bmatrix}$$

Thus

$$\begin{aligned} J_{\min} &= r_0 - \frac{r_1^2}{r_0^2 - r_1^2}[1, a] \begin{bmatrix} r_0 & -r_1 \\ -r_1 & r_0 \end{bmatrix} \begin{bmatrix} 1 \\ a \end{bmatrix} \\ &= r_0 - \frac{r_1^2}{r_0^2 - r_1^2}\left(r_0 - ar_1 + a^2 r_0 - ar_1\right) \end{aligned}$$

Now evaluate

$$\begin{aligned} r_0 &= ar_1 + ab\sigma_v^2 + \sigma_v^2(1 + b^2) \\ r_1 &= ar_0 + b\sigma_v^2 \end{aligned}$$

or

$$\begin{aligned} r_0 &= \frac{r_1}{2} + \frac{1}{2} + 2 \\ 2r_1 &= r_0 + 2 \end{aligned}$$

$r_1 = 3$, $r_0 = 4$.

First order, $J_{\min} = r_0 - 9/4$, second order $J_{\min} = r_0 - 18/7$.

Second order does better because $b \neq 0$.

Comment: This was a very popular question. Part (a) was answered well by most although many failed to properly account for the autocorrelation as some cross terms were dismissed as having no contribution. Part (b) very well answered. Part (c) was answered less well than expected. Errors in Part (a) meant many failed to get the right answer. Some were unable to correctly apply the solution of Part (b) to the calculation for the application to prediction.

Question 3

Answers:

a) A good discussion should include the following: the definition of the periodogram, the expression for its expected value and its variance, and the various improvements such as that of Bartlett, Blackman-Tukey and Welch, including the trade-off they offer.

b) $N\widehat{S}_X(e^{j\omega})$ is

$$\begin{aligned} &\left|\sum_{n=0}^{N-1} x_n e^{-j\omega n}\right|^2 \\ &= \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} x_m x_n e^{j\omega(m-n)} \\ &= \sum_{n=0}^{1} x_0 x_n e^{j\omega(0-n)} + \sum_{n=N-2}^{N-1} x_{N-1} x_n e^{j\omega(N-1-n)} + \sum_{m=1}^{N-2}\sum_{n=m-1}^{m+1} x_m x_n e^{j\omega(m-n)} + rem \end{aligned}$$

where rem denotes all other remaining terms. Note that *rem* has zero expectation. Taking the expectation gives

$$E\left\{\left|\sum_{n=0}^{N-1} x_n e^{-j\omega n}\right|^2\right\}$$
$$= \left(1+\alpha e^{-j\omega}\right) + \left(1+\alpha e^{j\omega}\right) + (N-2)\left(1+\alpha e^{j\omega}+\alpha e^{-j\omega}\right)$$
$$= N + (N-1)(\alpha e^{j\omega} + \alpha e^{-j\omega})$$

The true value is

$$S_X(e^{j\omega}) = \alpha e^{j\omega} + 1 + \alpha e^{-j\omega}.$$

So $\widehat{S}_X$ is biased for finite N but not for infinite N.

For

$$x_n = w_n + w_{n-1},$$

$R_{XX}[0] = 1$, $R_{XX}[-1] = R_{XX}[1] = 0.5$, other values of R_{XX} are zero. This is a special case with $\alpha = 0.5$.

Comments: Very popular question. Part (a) was answered well by most. Part (b) should have been a simple expansion of the absolute value followed by the application of the expectation operator but many could not execute it properly. Again those who found Part (b) difficult performed poorly in Part (c).

Question 4

Answers:

a) When stationary, x_0 will be a zero mean Gaussian.

$$E(x_n^2) = E(a^2 x_{n-1}^2 + w_n^2 + 2a x_{n-1} w_n)$$
$$= a^2 E(x_{n-1}^2) + E(w_n^2)$$

Thus the variance is $\sigma^2/(1-a^2)$.

The probability density of $(x_0, \ldots, x_n)$ is

$$p(x_0) \times \prod_{i=1}^n p(x_i|x_{i-1})$$

where $p(x_i|x_{i-1})$ is Gaussian with mean $a x_{i-1}$ and variance σ^2.

b) Including $p(x_0)$ will complicate the maximisation step. So maximise $p(x_1, \ldots, x_n|x_0)$ instead.

$\log p(x_1, \ldots, x_n|x_0)$ is

$$\log \frac{1}{(2\pi)^{n/2}} - \log \sigma^n + \frac{-1}{2\sigma^2} \sum_{i=1}^n (x_i - a x_{i-1})^2$$

Fix σ, differentiate w.r.t. a and set to zero to get

$$\frac{-1}{2\sigma^2} \sum_{i=1}^n -2(x_i - a x_{i-1}) x_{i-1} = 0$$

or

$$\widehat{a} = \frac{\sum_{i=1}^{n} x_i x_{i-1}}{\sum_{i=1}^{n} x_{i-1}^2}.$$

Use this value of a and solve for σ:

$$-\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i - \widehat{a}x_{i-1})^2 = 0$$

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \widehat{a}x_{i-1})^2$$

c) We see that

$$\widehat{a} = \frac{\widehat{R}_{XX}[1]}{\widehat{R}_{XX}[0]}$$

and

$$\begin{aligned}
\sigma^2 &= \frac{1}{n}\sum_{i=1}^{n} x_i^2 + \widehat{a}^2 x_{i-1}^2 - 2x_i\widehat{a}x_{i-1} \\
&= \widehat{R}_{XX}[0] + \frac{\widehat{R}_{XX}[1]}{\widehat{R}_{XX}[0]^2}^2 \widehat{R}_{XX}[0] - 2\frac{\widehat{R}_{XX}[1]^2}{\widehat{R}_{XX}[0]} \\
&= \widehat{R}_{XX}[0] - \frac{\widehat{R}_{XX}[1]^2}{\widehat{R}_{XX}[0]}
\end{aligned}$$

Which is the same as the Yule-Walker estimate. Here we have regarded both $n^{-1}\sum_{i=1}^{n} x_i^2$ and $n^{-1}\sum_{i=1}^{n} x_{i-1}^2$ as $\widehat{R}_{XX}[0]$.

d) Write the AR(P) model in state-space form:

$$\mathbf{x}_n = \begin{bmatrix} x_n \\ \vdots \\ x_{n-P+1} \end{bmatrix} = \mathbf{A}\mathbf{x}_{n-1} + \begin{bmatrix} 1 \\ 0 \\ \vdots \end{bmatrix} w_n$$

When stationary, $p(x_0, \ldots, x_{P-1})$ is a Gaussian density with zero mean and constant variance, say $\mathbf{S}$. Computing the variance of both sides of the state-space equation and equating them gives

$$\mathbf{S} = \mathbf{A}\mathbf{S}\mathbf{A}^{\mathrm{T}} + \sigma^2\mathbf{b}\mathbf{b}^{\mathrm{T}}$$

where $\mathbf{b} = [1, 0, \ldots, 0]^{\mathrm{T}}$. Now this equation can be solved for all the elements of the square matrix $\mathbf{S}$.

It is much simpler to maximise $p(x_P, \ldots, x_n | x_0, \ldots, x_{P-1})$ when n is much larger than P as there will be little loss of performance (statistical efficiency) in doing so. Note that $p(x_P, \ldots, x_n | x_0, \ldots, x_{P-1})$ can be easily written down and maximised as was done for the AR(1) case.

Comments: Overwhelmingly the least popular question. Part (a) was answered well via the probability chain rule. Part (b) was a simple least squares calculation and was correctly done. Part (d) proved challenging although it should have been an easy mark earner since all that was needed was a description of the procedure.