2007 IIB 4F7 Prof S Godsill
sjg@eng.cam.ac.uk

???? ? May 2007 9 to 10.30??

Module 4F7

DIGITAL FILTERS AND SPECTRUM ESTIMATION

Worked Solutions.

STATIONERY REQUIREMENTS SPECIAL REQUIREMENTS

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

Version 1 - 7 Feb 2007

1 Consider the real valued signal

$$u(n) = \beta u(n-1) + w(n),$$

where $w(n)$ is an i.i.d. (independent and identically distributed) sequence with zero mean and variance equal to one. The aim is to learn the unknown value β.

(a) Describe, without mathematical detail, the principal methods for adaptive estimation of a parameter such as β above. Compare and contrast their performance and computational load. [25%]

(b) Assume the signal $u(n)$ is stationary and derive $E\{u(n)^2\}$. [15%]

(c) Describe how you would use the least mean-squares (LMS) algorithm to learn β. What is the minimum mean-squared error (MSE) at the optimal solution? [20%]

(d) Describe how you would use the recursive least squares (RLS) algorithm to solve the same problem and show that the RLS algorithm converges to the optimal solution when its parameter $\lambda = 1$. [40%]

2 The inverse of a symmetric positive definite matrix $\mathbf{R}$ can be expressed as

$$\mathbf{R}^{-1} = \mu \sum_{k=0}^{\infty} (\mathbf{I} - \mu \mathbf{R})^k$$

where μ is a small positive constant and $\mathbf{I}$ is the identity matrix.

(a) Using the decomposition $\mathbf{R} = \mathbf{Q}^{\mathrm{T}} \Lambda \mathbf{Q}$ where $\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{Q}\mathbf{Q}^{\mathrm{T}} = \mathbf{I}$ and Λ is diagonal, find the range of values for μ for which the sum exists. [35%]

(b) The global solution to a Wiener filtering problem is given by $\mathbf{h}_{\text{opt}} = \mathbf{R}^{-1}\mathbf{p}$, and $\mathbf{h}(n)$ is the solution obtained at step n of an iterative scheme. Using the association $\mathbf{h}(n) = \mu \sum_{k=0}^{n-1} (\mathbf{I} - \mu \mathbf{R})^k \mathbf{p}$, derive the Steepest Descent recursion. [30%]

(c) Write down the cost function $J(\mathbf{h})$ for the Wiener filtering problem and sketch the contour diagram of $J(\mathbf{h})$ when $\mathbf{R} = \begin{bmatrix} 1.1 & 0.5 \\ 0.5 & 1.1 \end{bmatrix}$. Make an approximate sketch of the evolution $\mathbf{h}(n)$ of the Steepest Descent recursion on the same contour plot when (i) $\mu = 0.01$ and (ii) $\mu = 4$. [35%]

SECTION Answers to Question 1

a) Assume the signal $u(n)$ is stationary and derive $E\{u(n)^2\}$. [20%]

Answer:

$$E\{u(n)^2\} = \beta^2 E\{u(n-1)^2\} + E\{w(n)^2\} + 2\beta E\{w(n)u(n-1)\}.$$

Now $E\{w(n)u(n-1)\} = 0$. This implies

$$E\{u(n)^2\} = \beta^2 E\{u(n-1)^2\} + E\{w(n)^2\}$$

or

$$E\{u(n)^2\} = \frac{1}{1-\beta^2}.$$

b) Describe how you would use the LMS algorithm to learn β what is the minimum MSE at the optimal solution. [30%]

Answer:

Find h that minimises

$$E\{(u(n)-hu(n-1))^2\}$$
$$=(\beta-h)^2E\{u(n-1)^2\}+E\{w(n)^2\}+2(\beta-h)E\{w(n)u(n-1)\}.$$

Since $E\{w(n)u(n-1)\}=0$, it is obvious that the minimum MSE is

$$E\{w(n)^2\}=1.$$

The LMS update rule is

$$h(n+1)=h(n)+\mu(u(n)-h(n)u(n-1))u(n-1),$$

which is obtained by taking the gradient of the cost function w.r.to h.

c) Describe how you would use the RLS algorithm to solve the same problem and show that the RLS algorithm converges to the optimal solution when $\lambda=1$. [50%]

Answer:

The RLS algorithm solves

$$h(n)=\arg\min_h \sum_{k=1}^n \lambda^{n-k}(u(k)-hu(k-1))^2$$

Differentiating this function w.r.to h gives

$$-2\sum_{k=1}^n \lambda^{n-k}(u(k)-hu(k-1))u(k-1).$$

Setting this function to zero and solving for h gives

$$h=\frac{\sum_{k=1}^n \lambda^{n-k}u(k)u(k-1)}{\sum_{k=1}^n \lambda^{n-k}u(k-1)u(k-1)}.$$

When $\lambda=1$,

$$h=\frac{n^{-1}\sum_{k=1}^n u(k)u(k-1)}{n^{-1}\sum_{k=1}^n u(k-1)u(k-1)}.$$

As n tends to infinity, the denominator tends to $E\{u(n)^2\}=(1-\beta^2)^{-1}$. The numerator is

$$n^{-1}\sum_{k=1}^n (\beta u(k-1)u(k-1)+w(k)u(k-1))$$

which tends to

$$E(\beta u(n-1)u(n-1))+E(w(n)u(n-1))=\frac{\beta}{1-\beta^2}$$

since $E(w(n)u(n-1))=0$.

SECTION Answers to Question 2

a) Using the decomposition $\mathbf{R} = \mathbf{Q}^{\mathrm{T}}\blacksquare\mathbf{Q}$ where $\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{Q}\mathbf{Q}^{\mathrm{T}} = \mathbf{I}$ and $\blacksquare$ is diagonal, find the range of values for μ for which the sum exists. [35%]

Answer: One has

$$\begin{aligned}(\mathbf{I} - \mu\mathbf{R}) &= \left(\mathbf{I} - \mu\mathbf{Q}^{\mathrm{T}}\blacksquare\mathbf{Q}\right) \\ &= \mathbf{Q}^{\mathrm{T}}\left(\mathbf{I} - \mu\blacksquare\right)\mathbf{Q}\end{aligned}$$

and clearly, using $\mathbf{Q}\mathbf{Q}^{\mathrm{T}} = \mathbf{I}$,

$$\begin{aligned}&(\mathbf{I} - \mu\mathbf{R})^k = \mathbf{Q}^{\mathrm{T}}\left(\mathbf{I} - \mu\blacksquare\right)^k\mathbf{Q} \\ &\Rightarrow \textstyle\sum_{k=0}^{\infty}(\mathbf{I} - \mu\mathbf{R})^k = \mathbf{Q}^{\mathrm{T}}\left(\sum_{k=0}^{\infty}(\mathbf{I} - \mu\blacksquare)^k\right)\mathbf{Q}.\end{aligned}$$

So the sum exists provided

$$\sum_{k=0}^{\infty}(1 - \mu\lambda_i)^k$$

exists. Which is true $|1 - \mu\lambda_i| < 1$ or

$$-1 < 1 - \mu\lambda_i < 1$$

or $2 > \mu\lambda_i > 0$. So

$$\frac{2}{\lambda_{\max}} > \mu > 0.$$

b) Using $\mathbf{h}(n) = \mu\sum_{k=0}^{n-1}(\mathbf{I} - \mu\mathbf{R})^k\mathbf{p}$, derive the Steepest Descent recursion. [30%]

Answer:

$$\begin{aligned}\mathbf{h}(n+1) &= \mu\sum_{k=0}^{n}(\mathbf{I} - \mu\mathbf{R})^k\mathbf{p} \\ &= \mu(\mathbf{I} - \mu\mathbf{R})^0\mathbf{p} + (\mathbf{I} - \mu\mathbf{R})\,\mu\sum_{k=1}^{n}(\mathbf{I} - \mu\mathbf{R})^{k-1}\mathbf{p} \\ &= \mu\mathbf{p} + (\mathbf{I} - \mu\mathbf{R})\,\mathbf{h}(n) \\ &= \mathbf{h}(n) + \mu\left(\mathbf{p} - \mathbf{R}\mathbf{h}(n)\right)\end{aligned}$$

c) Write down the cost function $J(\mathbf{h})$ for the Wiener filtering problem and sketch the contour diagram of $J(\mathbf{h})$ when $\mathbf{R} = \begin{bmatrix} 1.1 & 0.5 \\ 0.5 & 1.1 \end{bmatrix}$. Sketch the evolution $\mathbf{h}(n)$ of the

Steepest Descent recursion on the same contour plot when i) $\mu = 0.01$ and (ii) $\mu = 4$. (Rough sketches conveying the shape will suffice!) [35%]

Answer

The cost function is

$$J(\mathbf{h}) = J_{\min} + (\mathbf{h} - \mathbf{h}_{\text{opt}})^{\text{T}}\mathbf{R}(\mathbf{h} - \mathbf{h}_{\text{opt}}).$$

The contour diagram looks something like

Convergence of Steepest Descent is assured provided $\frac{2}{\lambda_{\max}} > \mu > 0$. For the given $\mathbf{R}$, $\lambda_{\max} = 1.6$ and $\lambda_{\min} = 0.6$. For stability we need $\mu < 1.25$.

So for $\mu = 0.01$ it converges and $\mu = 4$ it diverges. The plot of Steepest Descent may look like the solid curved arrow for $\mu = 0.01$ and the dashed arrow for $\mu = 4$. The arrow indicates the path of $\mathbf{h}(n)$ as a function of iteration. Note the convergence of $\mathbf{h}(n)$ for $\mu = 0.01$ and divergence for $\mu = 4$.

3 (a) Describe the periodogram method for power spectrum analysis, and explain the principal means of improving its performance. [35%]

Solution: Periodogram is obtained from the DTFT of $\{x_0, x_1, ..., x_{N-1}\}$:

$$\boxed{\begin{aligned} \hat{S}_X(e^{j\omega T}) &= \frac{1}{N}|X_w(e^{j\omega T})|^2 \\ X_w(e^{j\omega T}) &= \sum_{n=0}^{N-1} x_n e^{-jn\omega T} \end{aligned}} \quad (1)$$

which is known as the **Periodogram**.

Improve it by averaging over time and frequency, using Bartlett, Welch or Blackman-Tukey methods.

(b) Show that the variance of the periodogram estimate is approximately proportional to the square of the power spectrum for a *Gaussian random process*, i.e. show that

$$\text{var}(\hat{S}_X(e^{j\omega T}) \approx (S_X(e^{j\omega T}))^2$$

where $\hat{S}_X(e^{j\omega T})$ is the periodogram estimate and $S_X(e^{j\omega T})$ is the true power spectrum. [35%]

Solution:

This comes from the lecture notes as follows:

We can rewrite a stationary random process as a white noise process with power spectrum equal to 1 driving a linear filter:

$$X(e^{j\omega T}) = E(e^{j\omega T})H(e^{j\omega T})$$

The power spectrum of such a process is:

$$S_X(e^{j\omega T}) = 1.|H(e^{j\omega T})|^2$$

Now, define as usual windowed versions of E and X:

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

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

(TURN OVER for continuation of Question 3

Now, supposing we have the DTFT of $x_{w,n}$ and $e_{w,n}$ as $X_w(e^{j\omega T})$ and $E_w(e^{j\omega T})$ we have the approximate result:

$$X_w(e^{j\omega T}) \approx E_w(e^{j\omega T})H(e^{j\omega T})$$

Note that filtering $e_{w,n}$ with filter $H(e^{j\omega T})$ does not give $x_{w,n}$ exactly, since there will be filtering 'end effects' caused by the windowing operation on e_n. The approximation becomes exact as the window length becomes very large.

We then have:

$$1/N|X_w(e^{j\omega T})|^2 \approx 1/N|E_w(e^{j\omega T})|^2|H(e^{j\omega T})|^2$$

and hence:

$$\begin{aligned} var(1/N|X_w(e^{j\omega T})|^2) &\approx var(1/N|E_w(e^{j\omega T})|^2)(|H(e^{j\omega T})|^2)^2 \\ &= var(1/N|E_w(e^{j\omega T})|^2)(S_X(e^{j\omega T}))^2 \end{aligned}$$

But, E is white Gaussian noise, whose periodogram has variance equal to $\sigma^4 = 1$. Hence:

$$\mathrm{var}(1/N|X_w(e^{j\omega T})|^2) = \mathrm{var}(\hat{S}_X(e^{j\omega T}) \approx (S_X(e^{j\omega T}))^2$$

as required.

(c) In a spectrum analyser unit, power spectrum estimates are updated sequentially frame by frame through a long sequence of data, as follows:

$$\begin{aligned} &\hat{S}_X^{(1)}(e^{j\omega T}) = S_X^{(1)}(e^{j\omega T}) \\ &\hat{S}_X^{(n)}(e^{j\omega T}) = \alpha \hat{S}_X^{(n-1)}(e^{j\omega T}) + (1-\alpha) S_X^{(n)}(e^{j\omega T}),\ n = 2, 3, ..., \end{aligned}$$

where $\hat{S}_X^{(n)}(e^{j\omega T})$ is the estimate obtained at frame n in the data and $S_X^{(n)}(e^{j\omega T})$ is the periodogram estimate obtained from frame n. The nth frame of data is defined as the set of data points $[x_{(n-1)N}, ..., x_{nN-1}]$ where N (the framelength) is the number of data points for the periodogram. α is a positive parameter less than 1.

If the data within different frames can be considered independent of one another, determine the variance of the estimate $\hat{S}_X^{(n)}(e^{j\omega T})$ at frame n, assuming that the variance result in part (b) holds exactly.

Solution:

Variances of independent random variables add.

(cont.

So, if $V_n = var(\hat{S}_X^{(n)}(e^{j\omega T}))$ we have:

$$V_n = \alpha^2 V_{n-1} + (1-\alpha)^2 var(S_X^{(n)}(e^{j\omega T})) = \alpha^2 V_{n-1} + (1-\alpha)^2 (S_X(e^{j\omega T}))^2$$

In the steady state, $V_n = V_{n-1} = V$, so

$$V(1-\alpha^2) = (1-\alpha)^2 (S_X(e^{j\omega T}))^2$$

or

$$V = \frac{(1-\alpha)}{(1+\alpha)} (S_X(e^{j\omega T}))^2$$

This is an improvement compared to the basic periodogram, as in part (b). This is due to the averaging, or smoothing effect of the analyser. When α is small there is little smoothing and the variance is close to the periodoram's variance. With α close to 1 the variance is small owing to greater smoothing effect.

Discuss how the method could be used for analysis of non-stationary signals whose power spectrum changes slowly and smoothly with time. [30%]

For non-stationary signals can still achieve some variance reduction with an intermediate value of α, so that the estimate can slowly adapt to the changing spectral content of the signal.

4 A moving average (MA) process of order Q is expressed in the following form:

$$x_n = \sum_{q=0}^{Q} b_q w_{n-q}$$

where $\{w_n\}$ is a zero-mean white noise process having unity variance.

(a) Determine the power spectrum of the MA process and show that the autocorrelation function takes the following form:

$$R_{XX}[r] = \begin{cases} \sum_{q=r}^{Q} b_q b_{q-r} & \text{if } |r| \leq Q \\ 0 & \text{if } r > Q \end{cases}$$

[30%]

Solution:

Power spectrum is:

$$S_X(\exp(j\omega T)) = |\sum_{i=0}^{Q} b_i \exp(-ji\omega T)|^2$$

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

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

Substituting for x_n and x_{n+r} from the MA model equation gives:

$$\begin{aligned} R_{XX}[r] &= E\left[\sum_{s=0}^{Q} b_s w_{n-s} \sum_{q=0}^{Q} b_q w_{n+r-q}\right] \\ &= \sum_{q=0}^{Q}\sum_{s=0}^{Q} b_s b_q E[w_{n+r-q} w_{n-s}] \\ &= \sum_{q=0}^{Q} b_{q-r} b_q \\ &= \sum_{q=0}^{Q} b_{q-r} b_q \end{aligned}$$

as required.

(b) The bilateral z-transform for a sequence h_n is defined as

$$\mathscr{Z}\{h_n\} = \sum_{-\infty}^{+\infty} h_n z^{-n}$$

Show that the bilateral z-transform of $R_{XX}[r]$ is equal to $B(z)B(z^{-1})$, where

$$B(z) = \sum_{n=0}^{Q} b_n z^{-n}$$

[20%]

Solution:

$\sum_{q=0}^{Q} b_{q-r} b_q$ is the convolution of $\{b_q\}$ with $\{b_{-q}\}$, where $b_q = 0$ for $q < 0$ and $q > Q$. but z-transform of $\{b_{-q}\}$ is $B(z^{-1})$. Hence, by z-transform convolution theorem, z-transform of $\sum_{q=0}^{Q} b_{q-r} b_q$ is product of $B(z)$ with $B(z^{-1})$, as required.

(c) If the Q zeros of the polynomial $B(z^{-1})$ are located at n_i, for $i = 1, ..., Q$, show that $B(z)$ has zeros at positions $1/n_i$ for $i - 1, ..., Q$. [10%]

Solution: $B(n_i^{-1}) = 0$ by definition. Hence $B(z)|_{z=1/n_i} = 0$, and thus zeros of $B(z)$ are at $1/n_i$,...

(d) Use these results to explain the *spectral factorisation* method for estimation of the parameters of a MA model. State clearly any assumptions you make about the model. [20%]

Solution:

Measure autocorrelation function. Part (b) shows that this equals $B(z)B(z^{-1})$. Hence find the roots of $B(z)B(z^{-1})$. These will be arranged in the symmetrical pairs n_i and $1/n_i$, inside and outside the unit circle, for $i = 0, ..., Q$. Then *assume* a minimum phase MA filter, so choose n_i that lie within the unit circle as the required ones and reconstruct the filter polynomial from those zeros.

(e) Two values of the autocorrelation function for a MA process are measured as follows:

$$R_{XX}[0] = 5, R_{XX}[1] = 2$$

Determine from this data a suitable MA model of order $Q = 1$. [20%]

Solution:

Note: $R_{XX}[-1] = R_{XX}[1]$. Need roots of

$$2z^{-1} + 5 + 2z$$

Solving, get $-2, -0.5$. Hence required root is -0.5 (inside unit circle).

Thus polynomial is:

$$B(z) \propto 1 + 0.5z^{-1}$$

But $R_{XX}[0] = \sum b_i^2$

Therefore filter is $b_0 = 4 \times 1 = 4$, $b_1 = 4 \times 0.5 = 2$.

END OF PAPER