ENGINEERING TRIPOS PART IIA

Thursday 12 May 2012 9.00 to 10.30

Module 3F3 - worked solutions

SIGNAL AND PATTERN PROCESSING

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

All questions carry the same number of marks.

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

There are no attachments.

STATIONERY REQUIREMENTS
Single-sided script paper

SPECIAL REQUIREMENTS
Engineering Data Book
CUED approved calculator allowed

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

1 (a) A digital filter with impulse response $\{h_n\}$ is to be implemented using fixed point arithmetic. Explain the issues that may arise when compared with a floating point implementation, and methods that may be be employed to reduce these issues. Your explanation should include an explanation of overflow, l_1 or l_2 scaling and frequency response scaling. [30%]

Solution: In fixed point arithmetic the precision of results is lower than those of floating point calculations using the same number of bits. Hence calculations will have greater errors, poorer SNR at the output, and detailed analysis of the performance may be harder. Even worse, it is possible for the results of calculations to overflow the range of the fixed point arithmetic and this would lead to clipping, or wrap-around effects in the output. These overflow problems can be alleviated or removed by scaling the outputs from each stage of the filter such that most, or no inputs, can generate overflow at the output at any stage of the filter. Examples of this are:

- l_1 scaling ensures that no signal can possibly overflow at each stage of the filtering. The scaling required is the l_1 sum of the impulse response $\sum_{n=0}^{\infty} |h_n|$, assuming that inputs are in the range -1 to 1.
- l_2 scaling is less severe and does not guarantee no overflow for all signals. The scaling required is $\sum_{n=0}^{\infty} |h_n|^2$.
- Frequency response scaling. In this version the maximum amplitude sinusoid is prevented from overflowing the system, It requires computation of the maximum of the system's frequency response. Again, it doesn't guarantee that arbitrary signals will not overload.

(b) A digital filter has the transfer function

$$H(z) = \frac{(1-0.1z^{-1})}{(1-0.9z^{-1})}$$

Determine the locations of any poles and zeros for the filter and sketch its frequency magnitude response over the range $\omega T = 0$ to 2π, marking on the frequencies and magnitudes where any maxima and minima occur. [20%]

Answer:

1 Pole at $z = 0.9$. One zero at $z = 0.1$.

You should hence expect a max/min at $\Omega = 0$ and π. In fact, the gain is at a maximum of 9 at $\Omega = 0$, since the pole is closer to the point $z = 1$ on the unit circle than the zero, reducing to a minimum of $1.1/1.9 = 0.56$ at $\Omega = \pi$.

(cont.

The figure 1 below shows the resulting magnitude plot

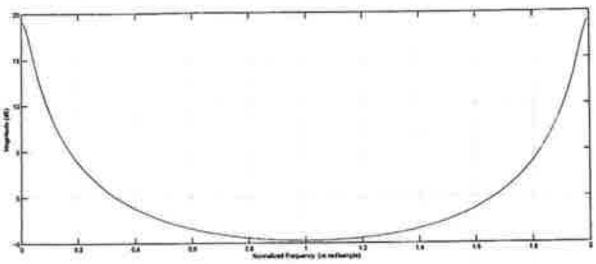

Fig. 1

(c) If $H^{(p)}(z)$ is the all-pole part of the transfer function of the above filter, i.e.

$$H^{(p)}(z) = \frac{1}{(1-z^{-1})},$$

and $h_n^{(p)}$ is the corresponding all-pole impulse response, determine the following quantities:

(i) $\sum_{n=0}^{\infty} |h_n^{(p)}|$
Answer:

$$H^{(p)}(z) = \frac{1}{(1-0.9z^{-1})}$$

So, by inverse z-transform:

$$h_n^{(p)} = 0.9^n$$

and, using sum of GP:

$$\sum_{n=0}^{\infty} |h_n^{(p)}| = \sum_{n=0}^{\infty} 0.9^n = 1/(1-0.9) = 10$$

(ii) $\sum_{n=0}^{\infty} |h_n^{(p)}|^2$ [20%]
Answer:

(TURN OVER for continuation of Question 1

Similarly:

$$\sum_{n=0}^{\infty} |h_n^{(p)}|^2 = \sum_{n=0}^{\infty} 0.81^n = 1/(1-0.81) = 5.26$$

Hence or otherwise determine a Direct Form II implementation of the original filter $H(z)$ that will reduce or eliminate the effects of overflow using l_1 scaling. You should assume that the maximum/minimum signal value which may be stored at input, output and any intermediate stages of the filter is ± 1, and your implementation should also make maximum use of any available head-room at any stage of the filter in order to give best quantisation noise performance. Explain whether overflow is completely prevented or just reduced by your implementation. [30%]

Answer:

l_1:

First consider the all-pole Part. The maximum signal output from this is given by the l_1 gain in part b) above, so an input scaling of 1/10 is required in order to prevent overflow at the output of the recursive part.

Then consider the overall response input to output:

$$h_n = h_n^{(p)} - 0.1h_{n-1}^{(p)} = [1, , (0.9-0.1), (0.9^2 - 0.1 * 0.9), (0.9^3 - 0.1 * 0.9^2)]$$

So,

$$\sum_{n=0}^{\infty} |h_n| = 1 + \sum_{n=0}^{\infty} (0.9-0.1)0.9^n = 1 + 0.8/(1-0.9) = 9$$

Hence we can apply a additional gain of 10/9=1.11 to the non-recursive coefficients of the filter, which will eliminate all possible overflows while achieving maximum signal-to-quantisation-noise ratio at the output.

The resulting implementation is:

$$H^{l_1}(z) = 0.1\frac{(1.11 - 0.111z^{-1})}{(1-0.9z^{-1})}$$

with an overall gain change of $0.1 * 1.11 = 0.111$ to the filter's original response.

l_1 completely eliminates the effects of overflow.

2 Examiner's comments:

The final parts caused problems for some - typically people didn't spot the simple way to calculate the impulse response using ergodicity and b)(i), and got the wrong idea about how one might compute the impulse response in (d). Only 1 candidate spotted how to compute the impulse response in (d)(ii).

(a) Define the terms white noise, wide-sense stationarity and ergodicity for a discrete time random process. Explain how ergodicity might enable a scientist to measure practically the autocorrelation function of a stationary random process occurring in an experiment. [30%]

Answer:

White noise is a stationary random process whose autocovariance function satisfies:

$$c[k] = \sigma^2 \delta_k$$

A stationary random process has constant statistical characteristics over time. In order to be wide-sense stationary we need a constant mean $E[X_t] = \mu$, constant autocorrelation function $r[k] = E[X_n X_{n+k}]$ (i.e. a function of time difference k only, and finite variance $E[(X_n - \mu)^2] < \infty$.

Ergodicity is a property in which a process 'forgets' its initial value given sufficient time. This means that expectations can be calculated from one member of the ensemble rather than from the whole ensemble.

If a process is correlation ergodic, then the autocorrelation function may be estimated as:

$$r[k] \approx \frac{1}{2N+1} \sum_{n=-N}^{+N} x_n x_{n+k}$$

for N 'sufficiently' large.

(b) It is proposed to measure the acoustical impulse response of a concert hall by playing zero-mean white noise through a loud-speaker on stage and recording (in digital form, sampled with appropriate anti-aliasing filters) the resulting sound at a key listening point in the hall.

If the white noise process is denoted by $\{W_n\}$ and that of the received sound at a key location is denoted by $\{X_n\}$, show that:

(i) $r_{WX}[k] = h_k \sigma_W^2$;

(ii) $r_{XX}[k] = \sigma_W^2 \sum_{i=0}^{\infty} h_i h_{k+i}$,

where σ_W is the standard deviation of the white noise and h_k, $k = 0, 1, 2, \ldots$, is the response of the hall to a digital impulse emitted by the loud-speaker at time $k = 0$ and recorded at the key location. [30%]

Answer:

$$
\begin{aligned}
r_{WX}[k] &= E[W_n \sum_{m=0}^{\infty} h_m W_{n-m+k}] \\
&= \sum_{m=0}^{\infty} h_m E[W_n W_{n-m+k}] = \sum_{m=0}^{\infty} h_m \sigma_W^2 \delta_{k-m} \\
&= h_k \sigma_W^2
\end{aligned}
$$

$$
\begin{aligned}
r_{XX}[k] &= E[\sum_{p=0}^{\infty} h_p W_{n-p}] \sum_{m=0}^{\infty} h_m W_{n-m+k}] \\
&= \sum_{p=0}^{\infty} \sum_{m=0}^{\infty} h_p h_m E[W_{n-p} W_{n-m+k}] \\
&= \sum_{p=0}^{\infty} \sum_{m=0}^{\infty} h_p h_m \sigma_W^2 \delta_{k-m+p} \\
&= \sum_{p=0}^{\infty} \sum_{m=0}^{\infty} h_p h_m \sigma_W^2 \delta_{k-m+p} \\
&= \sigma_W^2 \sum_{p=0}^{\infty} h_p h_{k+p}
\end{aligned}
$$

(c) Hence suggest a simple method for estimating the digital impulse response of the hall based purely on the white noise-generated recordings. [20%]

Answer:

Measure (estimate) the cross-correlation function between the input white noise and the measured signal X_n using the ergodicity properties of the white noise, and compute:

$$\hat{h}_k = \hat{r}_{WX}[k] / \sigma_W^2$$

(d) It is suggested that the process could be i) speeded up, or ii) made less intrusive by

(i) setting off a loud percussive noise, such as a gun, on stage and measuring the response at the key location;

(ii) playing recorded music through the loudspeaker and measuring the signal at the key location over a period of time.

Discuss whether recorded data from each of these two approaches might be analysed to obtain estimates of the hall's digital impulse response. What would be the pitfalls/advantages of these two methods compared with the original white noise approach?

Answer:

A gunshot would give an approximation to a delta-function at the stage. Hence measurement at a microphone of the time response would give an estimate of the impulse response of the hall. However, would be subject to environmental noise and could need to average over several (carefully time-aligned!) recorded responses.

Playing music would give a possible solution, provided the music is recorded at the loudspeaker output and synchronised with the microphone recording. Then, in the frequency domain, we could compute:

$$H(\Omega) = X(\Omega)/M(\Omega)$$

where $M(\Omega)$ is the (discrete) Fourier transform of the music and $X(\Omega)$ is the transform of the recorded microphone signal. This approach could be viable, but it would require synchronised DFTs from many frames of music data. Also, problems when there is a null in $M(\Omega)$, which would lead to infinities in the estimated H. Once again, some kind of averaging or regularisation over many time frames would probably be required, and it might not be possible to reconstruct from H a sufficiently long duration time response (concert halls resonate for many seconds typically).

[20%]

3 (a) Briefly describe the circumstances in which i) a matched filter, and ii) a Wiener filter, would be the appropriate tool for detection or estimation of a signal buried in additive white noise. Your description should identify any stationarity requirements for the signals and noise components for each type of filter. You should also identify a typical real-world application in each case. [30%]

Answer: i) A matched filter detects the presence of a known deterministic waveform in additive noise, which is assumed to be WSS. Examples include radar/sonar/demodulation of digital comms.. Wiener filters estimate random WSS signals in additive random WSS noise. Examples include noise reduction in telephony/audio systems.

(b) In a chemical experiment it is required to detect the presence or otherwise of a pulse waveform $p[t-L]$ buried in additive noise, where L is an unknown time offset for the pulse. The pulse waveform has the formula:

$$p[t] = \begin{cases} \exp(-\lambda t) & 0 \leq t \leq 20 \\ 0, & \text{otherwise} \end{cases}$$

where λ is a known constant.

Determine the coefficients of an optimal FIR filter which can detect the most likely time offset L for the pulse when it is observed in white noise v_t with zero mean and standard deviation σ_v:

$$x_t = p[t-L] + v_t$$

[you do not need to derive the form of this filter from first principles]. [20%]

Answer: Matched filter would be appropriate here. The FIR coefficients would be:

$$h_k = p[20-k], k = 0, \ldots, 20$$

(the time shift of 20 makes the filter causal, but otherwise introduces only a delay into the process).

(c) For the FIR filter in part b), determine:

(i) the mean-squared value of the output from the filter when just noise is present, i.e. $x_t = v_t$; [15%]

Answer:

$$E[y_t^2] = E[(\sum_{k=0}^{20} h_k v_{t-k})^2] = E[(\sum_{k=0}^{20} h_k v_{t-k})(\sum_{l=0}^{20} h_l v_{t-l})]$$
$$= E[(\sum_{k=0}^{20}\sum_{l=0}^{20} h_l h_k E[v_{t-k} v_{t-l}]]$$
$$= (\sum_{k=0}^{20}\sum_{l=0}^{20} h_l h_k \delta_{l-k} \sigma_v^2$$
$$= \sigma_v^2 \sum_{k=0}^{20} h_k^2 = \sigma_v^2 \sum_{k=0}^{20} \exp(-2\lambda k) = \sigma_v^2(1-\exp(-42\lambda))/(1-\exp(-2\lambda))$$

(ii) the maximum output from the filter when just the pulse is present, i.e. $x_t = p[t-L]$. [15%]

Answer:

Maximum output is when h_k is perfectly aligned with $p[t-L]$, then output from filter is:

$$\sum_{k=0}^{20} h_k p_k = \sum k = 0^{20} p_k^2 = (1-\exp(-42\lambda))/(1-\exp(-2\lambda))$$

Hence determine the maximum expected signal-to-noise ratio at the output of the filter.

An alternative simpler approach to detection of the pulse consists in finding the time location of the maximum value of the signal x_t from part (b) and using this as an estimate of L. Determine how much better the optimal filter will be in terms of maximum expected signal to noise ratio compared with this simple scheme. In particular, comment on performance as $\lambda \to 0$ and $\lambda \to \infty$. [20%]

Answer: Maximum signal-to-noise ratio occurs when output from signal component is at a maximum, in which case, SNR is (from previous part):

$$(\sum_{k=0}^{20} h_k p_k)^2/(\sigma_v^2(1-\exp(-42\lambda))/(1-\exp(-2\lambda))) = (1-\exp(-42\lambda))/(1-\exp(-2\lambda))/\sigma_v^2$$

[Note that we are assuming here that noise is small enough not to move the location of the maximum in the filter output].

Compared with the simple scheme, maximum value occurs there when $p[t-L]$ is at a maximum, i.e. at $e^{-\lambda 0} = 1$. In that case SNR is:

$$1/\sigma_v^2$$

[Note again that we are assuming here that noise is small enough not to move the location of the maximum x].

Compared to the matched filter, we have an improvement of

$$(1-\exp(-42\lambda))/(1-\exp(-2\lambda))$$

This ratio is always greater than 1, since $\exp(-42\lambda) << \exp(-2\lambda)$. How much greater than 1 will depend on λ. For small λ, the term $(1-\exp(-2\lambda))$ can be very close to zero, while the top line term is still close to 1, hence giving very large improvements potentially compared with the simple maximum detector.

Consider Taylor expansion for small λ:

$$(1-\exp(-42\lambda))/(1-\exp(-2\lambda)) = (1-(1-42\lambda...))/(1-(1-2\lambda...) \approx 42\lambda/2\lambda = 21$$

i.e. a possible improvement of 21, which would correspond to $p[t]$ being the rectangle pulse - in this case much more information can be gathered with the matched filter. With large values of λ we get no improvement, since the ratio equals 1. This also makes sense since then all the pulse energy is concentrated on $p[0]$ and no benefit would be gained from matched filtering.

4 Consider a dataset of observations $\mathscr{D} = \{(\mathbf{x}_n, y_n)\}$ where $n = 1, \ldots, N$, and N is the total number of data points. $\mathbf{x}_n$ is a two dimensional vector. A regression model of the following form is to be trained using the following form of regression

$$y_n = \mathbf{a}^{\mathrm{T}}\mathbf{x}_n + \varepsilon_n$$

where ε is independent zero-mean Gaussian noise with variance σ^2.

(a) Write down the log-likelihood $\log(p(y_1, \ldots, y_N | \mathbf{x}_1, \ldots, \mathbf{x}_N, \mathbf{a}, \sigma^2))$ in terms of $y_1, \ldots, y_N, \mathbf{x}_1, \ldots, \mathbf{x}_N, \mathbf{a}, \sigma^2$. [20%]

(b) Show that the maximum likelihood estimate of the regression parameters, $\hat{\mathbf{a}}$, can be expressed in the following form

$$\hat{\mathbf{a}} = \mathbf{C}^{-1}\mathbf{B}$$

You should clearly state the forms of the two matrices $\mathbf{C}$ and $\mathbf{B}$. The following equality may be useful

$$\frac{\partial \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = \mathbf{x}^{\mathrm{T}}\mathbf{A} + \mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$$

for any square matrix $\mathbf{A}$ and vector $\mathbf{x}$. [50%]

(c) A non-linear transformation $\phi(\mathbf{x}_n)$ is applied to the observations $\mathbf{x}_n$. The size of the resulting vector $\phi(\mathbf{x}_n)$ is d. Regression based on these transformed data points is then performed. Now

$$y_n = \mathbf{a}^{\mathrm{T}}\phi(\mathbf{x}_n) + \varepsilon_n$$

where ε is again independent zero-mean Gaussian noise with variance σ^2. Briefly discuss how the performance of the regression process and the estimation of the regression parameters may be impacted as the size of the transformed features, d, increases. [30%]

Solution:

5 *Regression and Maximum Likelihood estimation*

This question is fully covered in lectures.

(a) The log-likelihood can be written as

$$\begin{aligned}
\mathscr{L}(\mathbf{a}) &= \log(p(y_1,\ldots,y_N|\mathbf{x}_1,\ldots,\mathbf{x}_N,\mathbf{a},\sigma^2)) \\
&= \sum_{n=1}^{N} \log\left(\mathcal{N}(y_n - \mathbf{a}^{\mathrm{T}}\mathbf{x}_n; 0, \sigma^2)\right) \\
&= \sum_{n=1}^{N} \left(-\frac{1}{2}\log(\sqrt{2\pi\sigma^2}) - \frac{1}{2\sigma^2}\left[y_n^2 - 2y_n\mathbf{x}^{\mathrm{T}}\mathbf{a} + \mathbf{a}^{\mathrm{T}}\mathbf{x}\mathbf{x}^{\mathrm{T}}\mathbf{a}\right]\right)
\end{aligned}$$

(b) Rearranging the above expression

$$\mathscr{L}(\mathbf{a}) = \frac{N}{2}\log(\sqrt{2\pi\sigma^2}) - \frac{1}{2\sigma^2}\left(\left[\sum_{n=1}^{N} y_n^2\right] - 2\left[\sum_{n=1}^{N} y_n\mathbf{x}^{\mathrm{T}}\right]\mathbf{a} + \mathbf{a}^{\mathrm{T}}\left[\sum_{n=1}^{N}\mathbf{x}\mathbf{x}^{\mathrm{T}}\right]\mathbf{a}\right)$$

Differentiate this expression with respect to $\mathbf{a}$

$$\frac{\partial\mathscr{L}(\mathbf{a})}{\partial\mathbf{a}} = -\frac{1}{2\sigma^2}\left(-2\left[\sum_{n=1}^{N} y_n\mathbf{x}\right] + 2\left[\sum_{n=1}^{N}\mathbf{x}\mathbf{x}^{\mathrm{T}}\right]\mathbf{a}\right)$$

Equating this to zero and yields

$$\hat{\mathbf{a}} = \left(\sum_{n=1}^{N}\mathbf{x}\mathbf{x}^{\mathrm{T}}\right)^{-1}\left[\sum_{n=1}^{N} y_n\mathbf{x}\right]$$

Thus

$$\mathbf{C} = \sum_{n=1}^{N}\mathbf{x}\mathbf{x}^{\mathrm{T}}; \ \mathbf{D} = \sum_{n=1}^{N} y_n\mathbf{x}^{\mathrm{T}}$$

(c) Transforming the observation into a high-dimensional space $\phi(\mathbf{x}_n)$ should improve performance of modelling the training data as there are more parameters. Should mention:

- though better on the training data there may be issues generalising to unseen data

- the estimation requires an inversion of $\mathbf{C}$. As the dimensionality of the transformed data increases this may not be invertible due to the number of dimensions being greater than N, or numerical accuracy issues.

END OF PAPER

Version 1

3F3 Answers:

1.b) Pole at 0.9, zero at 0.1

2. c)(i) 10, (ii) 5.26