Delft University of Technology
Faculty of Electrical Engineering, Mathematics, and Computer Science
Circuits and Systems Group

EE4C03 STATISTICAL DIGITAL SIGNAL PROCESSING AND MODELING

9 November 2018, 13:30–16:30

Open book exam: copies of the book by Hayes and the course slides allowed. No other materials allowed.

This exam has four questions (40 points)

Question 1 (10 points)

A discrete-time random process $x(n)$ is generated by

$$x(n) = \sum_{k=1}^{p} a(k)x(n-k) \; + \; w(n)$$

where $w(n)$ is a zero mean white noise process with variance σ_w^2. Another process, $y(n)$, is formed by adding noise to $x(n)$,

$$y(n) = x(n) + v(n)$$

where $v(n)$ is zero mean white noise with a variance of σ_v^2 that is uncorrelated with $w(n)$.

(a) What is the power spectrum $P_x(e^{j\omega})$ of $x(n)$.

(*Hint:* first derive z-transform expressions.)

(b) What is the power spectrum $P_y(e^{j\omega})$ of $y(n)$.

(c) What are the Yule-Walker equations for the autocorrelation sequence $r_x(k)$?

Write them in a matrix form showing clearly how $r_x(k)$ $(k = 0, \cdots, p)$ can be computed.

(d) Consider $p = 1$. Derive a closed-form expression for the autocorrelation sequence $r_x(k)$ of $x(n)$.

Solution

(a) Since $x(n)$ is the output of an all-pole filter driven by white noise, $x(n)$ is an AR(p) process with a power spectrum

$$P_x(z) = \frac{1}{A(z)} \cdot \frac{1}{A(z^{-1})} \cdot \sigma_w^2$$

where $A(z) = 1 - \sum_{k=1}^{p} a(k)z^{-k}$. Also,

$$P_x(e^{j\omega}) = \frac{\sigma_w^2}{|A(e^{j\omega})|^2}$$

where

$$A(e^{j\omega}) = 1 - \sum_{k=1}^{p} a(k)e^{-jk\omega}$$

(b) The process is a sum of two uncorrelated random processes. Hence

$$r_y(k) = r_x(k) + r_v(k)\,,$$

$$P_y(e^{j\omega}) = P_x(e^{j\omega}) + P_v(e^{j\omega}) = \frac{\sigma_w^2}{|A(e^{j\omega})|^2} + \sigma_v^2$$

(c) The Yule-Walker equations are

$$r(k) - \sum_{l=1}^{p} a(l) r(k-l) = \sigma_w^2 \delta(k)\,, \qquad k \geq 0$$

or

$$\begin{bmatrix} r(0) & r(1) & \cdots & r(p) \\ r(1) & r(0) & & r(p-1) \\ \vdots & & \ddots & \vdots \\ r(p) & r(p-1) & \cdots & r(0) \end{bmatrix} \begin{bmatrix} 1 \\ -a(1) \\ \vdots \\ -a(p) \end{bmatrix} = \begin{bmatrix} \sigma_w^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

This can be rewritten as (cf. Hayes p.112)

$$\begin{bmatrix} 1 & -a(1) & \cdots & -a(p) \\ -a(1) & 1 & & -a(p-1) \\ \vdots & & \ddots & \vdots \\ -a(p) & -a(p-1) & \cdots & 1 \end{bmatrix} \begin{bmatrix} r(0) \\ r(1) \\ \vdots \\ r(p) \end{bmatrix} = \begin{bmatrix} \sigma_w^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

The correlation sequence follows by inversion:

$$\begin{bmatrix} r(0) \\ r(1) \\ \vdots \\ r(p) \end{bmatrix} = \sigma_w^2 \begin{bmatrix} 1 & -a(1) & \cdots & -a(p) \\ -a(1) & 1 & & -a(p-1) \\ \vdots & & \ddots & \vdots \\ -a(p) & -a(p-1) & \cdots & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

(d) In this case, we have (with $a = a(1)$, we assume $|a| < 1$)

$$x(n) = ax(n-1) \; + \; w(n)$$

which is the output of a filter

$$H(z) = \frac{1}{1 - az^{-1}}\,, \qquad h(n) = a^n u(n)$$

hence

$$r_x(k) = \sigma_w^2 \delta(k) * h(n) * h(-n) = \frac{\sigma_w^2}{1-a^2} a^{|k|}$$

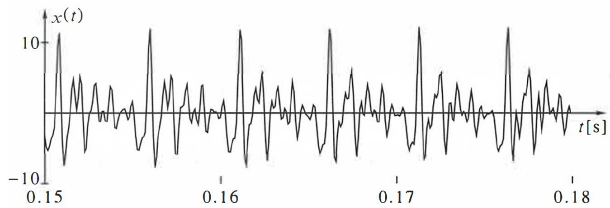
Alternative: take the inverse z-transform of

$$P_x(z) = \frac{\sigma_w^2}{(1-az^{-1})(1-az)}$$

See Hayes p.17 (table) for this.

Another alternative: write down the Yule-Walker equations and solve (cf. Hayes p.112)).

Question 2 (10 points)



The above signal is a 30 ms segment of voiced speech. Clearly it is quasi-periodic. The time between two pulses is the *pitch period* n_0. For the moment, we assume that n_0 is known.

We aim to model this signal as the output of an all-pole filter $H(z)$ (the vocal tract), driven by a pulse train of delta spikes (the vocal chords).

Let $x(n)$ be a subsequence of the signal consisting of two periods, length $N = 2n_0$, define the input sequence

$$v(n) = \delta(n) + \delta(n - n_0)$$

(two pulses spaced by n_0), and define the AR filter model

$$H(z) = \frac{b(0)}{1 + \sum_{k=1}^{p} a(k) z^{-k}} .$$

The signal model is $\hat{x}(n) = v(n) * h(n)$.

The Prony error is given by $\mathcal{E} = \sum_{n=0}^{N-1} e^2(n)$, where

$$e(n) = a(n) * x(n) - b(n) * v(n)$$

and $b(n) = b(0)\delta(n)$.

(a) Give a clear motivation for this choice of the Prony error.

(b) Derive the Prony normal equations that follow by minimizing the Prony error, resulting in conditions for $a(k)$ $(k = 1, \cdots, p)$ and $b(0)$.

(c) Give explicit (matrix) expressions from which we can directly compute $a(k)$ $(k = 1, \cdots, p)$ and $b(0)$ from the data.

(d) Now suppose the pitch period n_0 is not known. What is a simple and effective method to estimate n_0 from the data? (Note: deriving n_0 from the distance between the peaks of $x(n)$ is not always sufficiently accurate — what is a better method?)

Solution

(a) (The expression is a bit different from what is seen in the rest of the book. In particular, there is a term involving $b(0)$ which is often omitted in the book. Note the summation indices on $\mathcal{E}$.)

We would like to minimize $\sum [x(n) - \hat{x}(n)]^2$, but as shown in the book this leads to non-linear equations. Therefore, the trick is to minimize a filtered error, $\mathcal{E} = \sum [a(n) * \{x(n) - \hat{x}(n)\}]^2$. Since $a(n) * \hat{x}(n) = b(n) * v(n)$, the result follows.

(b) The normal equations follow by setting the derivative of $\mathcal{E}$ to $a(k)$ and $b(0)$ equal to zero. With $e(n) = \sum a(k)x(n-k) - b(0)\delta(n) - b(0)\delta(n-n_0)$, we find

$$\frac{\partial \mathcal{E}}{\partial a(k)} = \sum_{n=0}^{N-1} 2e(n)x(n-k) = 0\,, \qquad k = 1, \cdots, p$$

Inserting the expression for $e(n)$ and defining

$$r_x(k,l) = \sum_{n=0}^{N-1} x(n-l)x(n-k)$$

then these equations can be written as

$$\sum_{l=1}^{p} a(l)r_x(k,l) - b(0)x(-k) - b(0)x(n_0-k) = -r_x(k,0)\,, \qquad k = 1, \cdots, p$$

Similarly, we find

$$\frac{\partial \mathcal{E}}{\partial b(0)} = \sum_{n=0}^{N-1} 2e(n)[\delta(n) + \delta(n-n_0)] = 0$$

which can be written as

$$x(0) - b(0) + \sum_{l=1}^{p} a(l)x(n_0-l) - b(0) = -x(n_0)\,.$$

(c) We can collect the equations in vector/matrix form. Assume exact periodicity for $n < 0$: $x(n) = x(n_0 - n)$ for $n < 0$, and define

$$\begin{aligned} \mathbf{x} &= [x(n_0-1), x(n_0-2), \cdots, x(n_0-p)]^T \\ \mathbf{a} &= [a(1), \cdots, a(p)]^T \\ \mathbf{r}_x &= [r_x(1,0), \cdots, r_x(p,0)]^T \end{aligned}$$

and a $p \times p$ correlation matrix $\mathbf{R}_x$ as usual, then the normal equations for $\{a(k)\}$ are

$$\mathbf{R}_x\mathbf{a} - 2b(0)\mathbf{x} = -\mathbf{r}_x\,.$$

For $b(0)$ we find $\mathbf{x}^T\mathbf{a} - 2b(0) = -x(0) - x(n_0)$. Altogether,

$$\begin{bmatrix} \mathbf{R}_x & \mathbf{x} \\ \mathbf{x}^T & 1 \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ -2b(0) \end{bmatrix} = - \begin{bmatrix} \mathbf{r}_x \\ x(0) + x(n_0) \end{bmatrix}$$

(d) Form the autocorrelation sequence $r_x(k)$. There is a periodicity with period n_0, and $r_x(n_0)$ is nearly the same as $r_x(0)$. We can detect the distance between the peaks in $r_x(k)$; these would also be visible in the corresponding $P_x(\omega)$. This is more accurate than doing this on $x(n)$ as the correlation sequence involves all the data.

Using MUSIC is dubious and has to be motivated—is this a sum of sinusoids (harmonic signal)?

Question 3 (7 points)

Let $y(t)$ be a signal of the form

$$y(t) = \sin(2\pi f_1 t + \phi_1) \ + \ \sin(2\pi f_2 t + \phi_2) \ + \ w(t), \quad t \in [0, T]$$

where $w(t)$ is a zero mean white Gaussian noise process of unit variance and $\phi_1, \ \phi_2$ are unknown phase components. Also the frequencies f_1 and f_2 are unknown.

You are provided with T seconds of this signal and the task is to estimate its spectrum using the Bartlett method. The parameter that you can control is the sampling frequency, f_s.

(a) We want to attain the following objectives:

- The two frequencies f_1 and f_2 are sufficiently resolved.
- The variance of the spectral estimate is less than 1% of the square of the true spectral density.

Derive an expression for the lowest possible f_s that attains these objectives.

Now, assume that the noise variance σ_w^2 is unknown and that the sampling frequency f_s is given. Assume also that you have the option to choose between using the Periodogram method and using the Bartlett method.

(b) Which of these two alternatives do you recommend if you are primarily interested in estimating the noise level σ_w^2. Motivate your answer.

(c) Which of these two alternatives do you recommend if you are primarily interested in getting good estimates of f_1 and f_2, and that $f_2 - f_1$ may be very small. Motivate your answer.

Solution

(a) (First of all, we need to assume that the Nyquist condition is satisfied. But this condition doesn't actually play a role as we can allow aliasing if we first use a bandlimiting filter that selects only the interval in which f_1 and f_2 are present.)

Use Hayes Table 8.7: the variance is of the Bartlett estimate is $\frac{1}{K}[P_x(e^{j\omega})]^2$, and the resolution $\Delta\omega$ is $0.89K\frac{2\pi}{N}$. In this method, the total number of samples N is split into K blocks of L samples. We thus find

$$\frac{1}{K} \leq 0.01 \quad \Rightarrow \quad K \geq 100$$

The normalized frequency is

$$\omega = \Omega T = 2\pi \frac{f}{f_s} \qquad \Rightarrow \qquad \Delta\omega = 2\pi \frac{\Delta f}{f_s}$$

Thus (taking $K = 100$)

$$\frac{\Delta f}{f_s} = \frac{0.89K}{N} = \frac{89}{N} \leq \frac{|f_2 - f_1|}{f_s} \quad \Rightarrow \quad N \geq \frac{89\, f_s}{|f_2 - f_1|}$$

Using $N = Tf_s$, we find

$$T \geq \frac{89}{|f_2 - f_1|}$$

and f_s doesn't play a role! (This relates to the famous "time-bandwidth product".)

(b) Bartlett, because (1) unlike the Periodogram it is a consistent estimate that gives a small variance due to the averaging, and (2) windowing gives low sidelobes due to the two sinusoids.

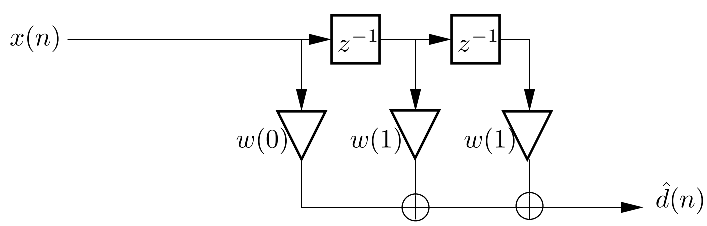
(c) Periodogram, because it gives the best resolution, and we are not interested in consistency of the power estimate here.

Question 4 (13 points)

Let $x(n)$ and $d(n)$ be two zero-mean real-valued random processes, with auto- and crosscorrelation coefficients given in the following table:

k	0	1	2	3	4	...
$r_d(k) = E\{d(n)d(n-k)\}$	1	0	0	0	0	...
$r_x(k) = E\{x(n)x(n-k)\}$	2	1	1	1	1	...
$r_{dx}(k) = E\{d(n)x(n-k)\}$	1	0	0	0	0	...

Assuming we have measured $x(n)$, we wish to estimate the unknown process $d(n)$ from $x(n)$ using the following FIR filtering structure:



The filter has three tap coefficients, but $w(1)$ and $w(2)$ are equal. The goal is to find the two filter coefficients $w(0)$ and $w(1)$ that minimize the mean square error (MSE)

$$\xi = E\{|d(n) - \hat{d}(n)|^2\}$$

(a) In view of the correlation table, mention a specific application of the resulting filter.

(b) Derive the Wiener-Hopf equations and their solution to determine the optimal filter coefficients $w(0)$ and $w(1)$, as well as the expression for the minimum MSE ξ.

(c) Assuming the values of $r_d(k)$, $r_x(k)$, and $r_{dx}(k)$ as given in the table, what are the values of these optimal filter coefficients and for the minimum MSE?

(d) In the above, we constrained the structure of the filter to satisfy $w(1) = w(2)$. Why does this make sense?

(Hint: check the equations for the unstructured case and compare.)

(e) Suppose we want to compute $w(0)$ and $w(1)$ in an adaptive fashion. Assume for the moment that we have a reference signal $d(n)$ available. Give the expression for the least mean squares (LMS) update equations for $w(0)$ and $w(1)$. Also give appropriate initial conditions and derive a good range for the step size.

(f) Similarly, give the expressions for the exponentialy weighted recursive least squares (RLS) update equations for $w(0)$ and $w(1)$. Also give appropriate initial conditions and a good range for the weighting factor.

(g) In practice, we do not have $d(n)$ available as a reference signal. How can we implement the adaptive algorithms in this case?

Solution

(a) The signal $d(n)$ has no correlations except at lag 0, it could be a very wideband signal, e.g. white noise.

Let us model the signal $x(n)$ as $x(n) = d(n)+v(n)$, where $v(n)$ is a disturbance independent from $d(n)$. Then $r_x(k) = r_d(k) + r_v(k)$, hence $r_v(k) = 1$ (constant). Hence $v(k)$ is fully correlated (deterministic), e.g. a very narrowband signal.

The purpose of the filter would be to filter out the disturbing signal $v(n)$. This is similar to the application mentioned in the book, where (reversely) the white noise signal is filtered from a sinusoid in noise.

(b) Defining the inputs of the filter $z_0(n) = x(n)$ and $z_1(n) = x(n-1) + x(n-2)$, we can write our filtering process as

$$\hat{d}(n) = w(0)z_0(n) + w(1)z_1(n) = \mathbf{w}^T\mathbf{z},$$

where $\mathbf{w} = [w(0), w(1)]^T$ and $\mathbf{z} = [z_0(n), z_1(n)]^T$. The Wiener-Hopf equations for the optimal filter can be written as

$$\mathbf{w}_o = \mathbf{R}_z^{-1}\mathbf{r}_{dz}$$

where

$$\mathbf{R}_z = E\{\mathbf{z}\mathbf{z}^T\} = \begin{bmatrix} r_x(0) & r_x(1) + r_x(2) \\ r_x(1) + r_x(2) & 2r_x(0) + 2r_x(1) \end{bmatrix}$$

$$\mathbf{r}_{dz} = E\{d(n)\mathbf{z}\} = \begin{bmatrix} r_{dx}(0) \\ r_{dx}(1) + r_{dx}(2) \end{bmatrix}$$

The expression for the minimal MSE is

$$\xi_o = r_d(0) - \mathbf{r}_{dz}^T\mathbf{w}_o = r_d(0) - \mathbf{r}_{dz}^T\mathbf{R}_z^{-1}\mathbf{r}_{dz}.$$

(c) For the values given in the table, we obtain $w_o(0) = 3/4$, $w_o(1) = -1/4$, and $\xi_o = 1/4$.

(d) If we remove the structure, we look for a filter of the form

$$\hat{d}(n) = w(0)x(n) + w(1)x(n-1) + w(2)x(n-2) = \mathbf{w}^T\mathbf{x},$$

where $\mathbf{w} = [w(0), w(1), w(2)]^T$ and $\mathbf{x} = [x(n), x(n-1), x(n-2)]^T$. The Wiener-Hopf equations and the minimal MSE are then respectively given by the standard expressions

$$\mathbf{w}_o = \mathbf{R}_x^{-1}\mathbf{r}_{dx}$$

and

$$\xi_o = r_d(0) - \mathbf{r}_{dx}^T\mathbf{w}_o = r_d(0) - \mathbf{r}_{dx}^T\mathbf{R}_x^{-1}\mathbf{r}_{dx}.$$

For the values in the table, this leads to $\mathbf{w}_o = [3/4, -1/4, -1/4]^T$ and $\xi_o = 1/4$, which is exactly the same result as in (b). This is because both the autocorrelation function $r_x(k)$ as well as the crosscorrelation function $r_{dx}(n)$ have the same values at lags 1 and 2.

(e) Standard expressions from the book and slides.

(f) Standard expressions from the book and slides.

(g) As in the application in the book (section 9.2.5, p.516), if we don't have a reference signal $d(n)$ we could take $d(n) = x(n-1)$. However, then the adaptive filter will converge to $v(n)$ and filter out $d(n)$ (since it is uncorrelated for lags larger than 0). Therefore, we should, as final step, subtract the output of the adaptive filter from $x(n)$.