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

EE4C03 STATISTICAL DIGITAL SIGNAL PROCESSING AND MODELING

4 November 2020, 13:30–16:30

Block 1 (13:30-15:00)

Open book, strictly timed take-home exam. (Electronic) copies of the book and the course slides allowed. No other tools except a basic pocket calculator permitted.

Upload answers during 14:55–15:05

This block consists of three questions (25 points); more than usual, and this will be taken into account during grading. Answer in Dutch or English. Make clear in your answer how you reach the final result; the road to the answer is very important. Write your name and student number on each sheet.

Hint: Avoid losing too much time on detailed calculations, write down the general approach first.

Question 1 (8 points)

(a) Let $x(n) = 4 + v(n)$, where $v(n)$ is real-valued zero mean i.i.d. noise with variance $\sigma_v^2 = 1$. The signal is filtered by an FIR filter with impulse response

$$h(n) = \begin{cases} 1 & n = 0, \\ -\frac{1}{2} & n = 1, \\ 0 & \text{otherwise.} \end{cases}$$

The output of the filter is $y(n) = h(n) * x(n)$.

Compute the mean value $m_y(n) = E\{y(n)\}$ and the autocovariance sequence $c_y(k)$.

(b) For a zero mean WSS process, let the autocorrelation sequence be given by

$$r(k) = A\cos(\omega_0 k)\,,$$

and construct the $N \times N$ correlation matrix $\mathbf{R}$.

Derive that the rank of $\mathbf{R}$ is equal to 2.

Hint: first write $\cos(\omega_0 k)$ *as the sum of two exponentials.*

(c) *True* or *False:* is this a valid autocorrelation matrix?

$$\mathbf{R} = \begin{bmatrix} 2 & 2 & 0 & 0 \\ 2 & 2 & 2 & 0 \\ 0 & 2 & 2 & 2 \\ 0 & 0 & 2 & 2 \end{bmatrix}$$

(Motivate your answer.)

(d) Suppose we have N samples of a zero mean WSS random process $x(n)$, but with one sample missing:

$$\mathbf{x} = [x(0), x(1), \cdots, x(n_0 - 1), x(n_0 + 1), \cdots, x(N-1)]^T$$

As usual, denote the correlation sequence by $r(k) = E\{x(n+k)x^*(n)\}$, and denote the correlation matrix $\mathbf{R} = E\{\mathbf{x}\mathbf{x}^H\}$.

Write down $\mathbf{R}$ in terms of $r(k)$. Is this a Toeplitz matrix? And can the usual correlation matrix corresponding to a sequence with no missing samples be reconstructed?

(e) *True* or *False:* if $y(n) = x_1(n) + x_2(n)$, and $x_1(n)$ is an AR(1) process and $x_2(n)$ is an AR(2) process, then $y(n)$ is an AR(3) process. (Motivate your answer.)

Question 2 (8 points)

Let

$$x(n) = \sin(\omega_0 n + \phi_0)\,,$$

where ϕ_0 is a random variable, uniformly distributed in the range $[-\pi, \pi]$, and ω_0 is fixed.

(a) Show that $x(n)$ is zero mean, and derive that the autocovariance sequence of $x(n)$ is

$$r_x(k) = \tfrac{1}{2}\cos(\omega_0 k)\,.$$

(b) Prove that $x(n)$ satisfies the identity

$$x(n) - 2\cos(\omega_0)x(n-1) + x(n-2) = 0$$

Hint: $\sin(A+B) = \sin(A)\cos(B) + \cos(A)\sin(B)$.

(c) Explain that this implies that $x(n)$ is "*deterministic*" (predictable).

(d) For a specific ω_0, the first few samples of $r_x(k)$ are

$$r_x(0) = 0.5\,, \qquad r_x(1) = 0.25\,, \qquad r_x(2) = -0.25\,, \qquad r_x(3) = -0.5 \qquad \cdots$$

Using the Schur algorithm, compute, step by step, the reflection coefficients Γ_1 and Γ_2.

(e) Looking at $|\Gamma_2|$, what can you conclude?

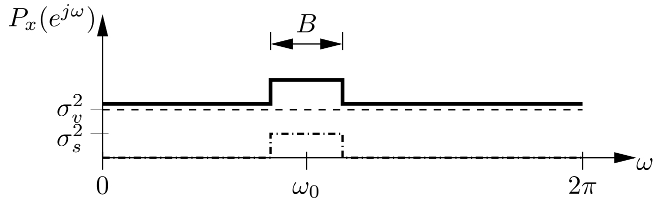
(f) Suppose now that $x(n)$ is the sum of 2 sinusoids. What can you say about the reflection coefficients?

Question 3 (9 points)

We are given 2000 samples of a single signal in noise,

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

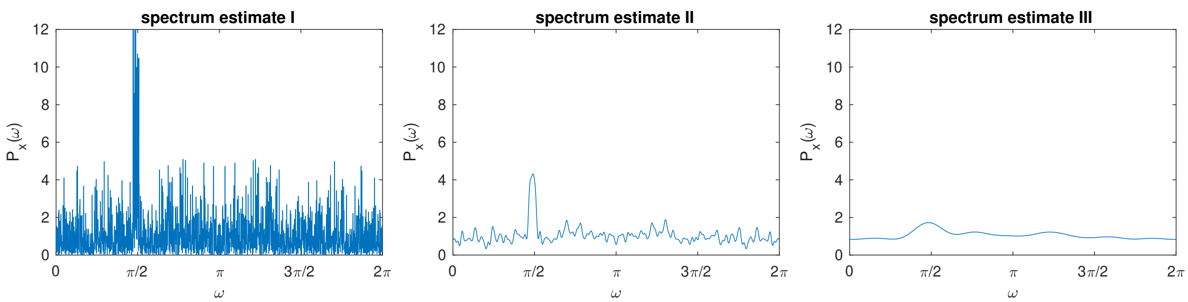
where $v(n)$ is zero mean white Gaussian noise with variance $\sigma_v^2 = 1$, while $s(n)$ is a weak astronomical signal, modeled as zero mean Gaussian noise with power spectrum $P_s(e^{j\omega})$, the dash-dot line in the figure. Based on physics, we expect the signal at $\omega_0 = 1.5$ rad and with a bandwidth $B = 0.16$ rad. We do not know its power in the passband.



We wish to detect the presence of the astronomical signal from a Bartlett spectrum estimate $\hat{P}_B(e^{j\omega})$. The $N = 2000$ samples are split into K blocks of L samples, $KL = 2000$. Three different options for (L, K) are considered:

$$(i) \quad L = 10, K = 200; \qquad (ii) \quad L = 100, K = 20; \qquad (iii) \quad L = 2000, K = 1.$$

Assume for the moment $\sigma_s^2 = 4$. The resulting spectrum estimates for $P_x(e^{j\omega})$, in random order, are:



(a) Which resolution is obtained for each choice of (L, K)? Also indicate how this relates to B.

(b) What is the variance on $\hat{P}_B(e^{j\omega})$ in each case? (Make a distinction between outside/inside the passband of $P_s(e^{j\omega})$.)

(c) Which spectrum estimate corresponds to which parameter set (L, K)?

(d) Explain the observed differences between spectrum I and II. In particular, explain why the height of the observed peak in spectrum II is much lower than that in spectrum I.

(e) Explain the observed differences between spectrum II and III. In particular, explain why the height of the observed peak in spectrum III is much lower than that in spectrum II.

(f) Which spectrum estimate would you use to estimate σ_s^2, and how would you estimate it? What is the variance of that estimate?

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

EE4C03 STATISTICAL DIGITAL SIGNAL PROCESSING AND MODELING

4 November 2020, 13:30–16:50
Block 2 (15:20-16:50)
Open book, strictly timed take-home exam. (Electronic) copies of the book and the course slides allowed. No other tools except a basic pocket calculator permitted.
Upload answers during 16:45–16:55

This block consists of three questions (25 points); more than usual, and this will be taken into account during grading. Answer in Dutch or English. Make clear in your answer how you reach the final result; the road to the answer is very important. Write your name and student number on each sheet.

Hint: Avoid losing too much time on detailed calculations, write down the general approach first.

Question 4 (9 points)

Consider a stochastic process consisting of a complex exponential in complex noise. Mathematically, this can be written as

$$x(n) = \alpha e^{j\omega_0 n} + \beta v(n), \tag{1}$$

where α and β are some weighting parameters, ω_0 is the frequency of the complex exponential and $v(n)$ is complex white Gaussian noise with mean 0 and variance 1.

We will fit an all-pole model of order two to this process, which can be written as

$$H(z) = \frac{b(0)}{1 + a(1)z^{-1} + a(2)z^{-2}},$$

with filter coefficients $b(0)$, $a(1)$ and $a(2)$.

Let us first assume that $\alpha = 1$ and $\beta = 0$, i.e., $x(n) = e^{j\omega_0 n}$, and that $N+1$ samples from $x(n)$ are given, i.e., $x(0), x(1), \ldots, x(N)$.

(a) Use the autocorrelation method based on these $N+1$ samples to compute $a(1)$ and $a(2)$. Provide all the steps of the derivation.

Hint: The inverse of the matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ *is given by* $\frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

(b) Use the covariance method based on these $N+1$ samples to compute $a(1)$ and $a(2)$. Provide all the steps of the derivation.

Let us now assume the general case where both α and β are non-zero.

(c) Compute the correlation sequence of $x(n)$ based on the model (1). Is the process $x(n)$ wide-sense stationary? Why or why not?

(d) Use the Yule Walker model to compute $a(1)$ and $a(2)$. Provide all steps.

(e) If you set $\alpha = 1$ and $\beta = 0$, how does the solution of (e) look like? Does it correspond to the solution of (a) and/or (b) above?

(f) If you now set $\alpha = 0$ and $\beta = 1$, how does the solution of (e) look like? Can you interpret that solution? Or in other words, is this the solution you would expect?

Question 5 (8 points)

In this question, we will develop the concept of optimal filtering but applied to compressed signals. Assume all signals are real-valued. Let us consider a stationary stochastic process $x(n)$ which depends on some stationary desired signal $d(n)$. The auto-correlation sequence of $x(n)$ is given by $r_x(k) = E\{x(n)x(n-k)\}$ and the cross-correlation sequence between $d(n)$ and $x(n)$ by $r_{dx}(k) = E\{d(n)x(n-k)\}$.

The signal $x(n)$ is now compressed using the following steps: i) $x(n)$ is split into non-overlapping windows of size N, denoted as $\mathbf{x}_i = [x(iN), x(iN+1), \ldots, x(iN+N-1)]^T$; ii) every window $\mathbf{x}_i$ is compressed by applying an $M \times N$ matrix $\mathbf{C}$ to it ($M < N$), leading to a smaller window $\mathbf{y}_i = \mathbf{C}\mathbf{x}_i$; iii) the smaller windows $\mathbf{y}_i$ are concatenated leading to the sequence $y(m)$. In short, every N samples of $x(n)$ are transformed into $M < N$ samples of $y(m)$ using the matrix $\mathbf{C}$.

(a) Express the auto-correlation matrix $\mathbf{R}_x = E\{\mathbf{x}_i\mathbf{x}_i^T\}$ as a function of $r_x(k)$.

Hint: Focus on a single entry of the matrix and compute the related correlation.

(b) If we define the windowed version of $d(n)$ as $\mathbf{d}_i = [d(iN), d(iN+1), \ldots, d(iN+N-1)]^T$, express the cross-correlation matrix $\mathbf{R}_{dx} = E\{\mathbf{d}_i\mathbf{x}_i^T\}$ as a function of $r_{dx}(k)$.

Hint: Use the same hint as in (a).

(c) Express the auto-correlation matrix $\mathbf{R}_y = E\{\mathbf{y}_i\mathbf{y}_i^T\}$ as a function of $\mathbf{R}_x$ from (a). Similarly, express the cross-correlation matrix $\mathbf{R}_{dy} = E\{\mathbf{d}_i\mathbf{y}_i^T\}$ as a function of $\mathbf{R}_{dx}$ from (b).

(d) Suppose we want to estimate $\mathbf{d}_i$ from $\mathbf{y}_i$ using an $M \times N$ filter $\mathbf{W}$ such that $\hat{\mathbf{d}}_i = \mathbf{W}^T\mathbf{x}_i$. Find the optimal filter $\mathbf{W}$ by solving

$$\min_{\mathbf{W}} E\{\|\mathbf{W}^T\mathbf{y}_i - \mathbf{d}_i\|^2\}.$$

Give all the steps of the derivation and express the solution as a function of $\mathbf{R}_x$ and $\mathbf{R}_{dx}$ using the results derived in (c).

Hint: Solve the problem for the kth filter $\mathbf{w}_k$ (kth column of $\mathbf{W}$) which is used to estimate $d(iN+k-1)$ (kth entry of $\mathbf{d}_i$). Then stack all the solutions.

(e) Suppose now that $x(n) = d(n)$ and both are white Gaussian noise with mean 0 and variance 1. How can then $\hat{\mathbf{d}}_i$ be expressed as a function of $\mathbf{d}_i$? Is $\hat{\mathbf{d}}_i = \mathbf{d}_i$? Why or why not?

(f) Based on your derivation in (d), express the steepest gradient descent recursion to compute $\mathbf{W}$. From that recursion, derive the LMS algorithm to adaptively compute $\mathbf{W}$ using knowledge of $\mathbf{y}_i$ and $\mathbf{d}_i$.

Question 6 (8 points)

Suppose we want to use spectral estimation techniques for localizing an acoustic source that is picked up by a number of microphones. The source is at position ϕ_0 (we simply use a scalar index here to indicate the position) and transmits a white Gaussian signal $s(n)$ with mean 0 and variance 1. Microphone k, with $k = 1, 2, \ldots, K$, is at position ϕ_k and picks up the signal

$$x_k(n) = a_k(\phi_0)s(n) + v_k(n),$$

where $a_k(\phi)$ is the attenuation from position ϕ to position ϕ_k, and $v_k(n)$ is spatially and temporally white Gaussian noise with mean 0 and variance σ^2. We also define the column vectors $\mathbf{x}(n) = [x_1(n), \ldots, x_K(n)]^T$, $\mathbf{v}(n) = [v_1(n), \ldots, v_K(n)]^T$, and $\mathbf{a}(\phi) = [a_1(\phi), \ldots, a_K(\phi)]^T$. Although not true in practice, we assume that $\mathbf{a}(\phi)$ is normalized for every position ϕ, i.e., $\|\mathbf{a}(\phi)\| = 1, \ \forall\phi$.

(a) Express $\mathbf{x}(n)$ as a function of $\mathbf{v}(n)$, $\mathbf{a}(\phi_0)$, and $s(n)$. Based on this model, compute the correlation matrix of $\mathbf{x}(n)$, denoted by $\mathbf{R}_x = E\{\mathbf{x}(n)\mathbf{x}^T(n)\}$.

(b) If the noise variance $\sigma^2 = 0$, then what is the rank of $\mathbf{R}_x$ and why?

We now want to estimate the source position ϕ_0 using spectral estimation techniques. Therefore, we first consider an arbitrary position ϕ and we apply the *position-dependent* filter $\mathbf{w}$ to $\mathbf{x}(n)$ leading to the output $y(n) = \mathbf{w}^T\mathbf{x}(n)$. The output power $P(\phi) = \mathbf{w}^T\mathbf{R}_x\mathbf{w}$, which can be interpreted as the "position" spectrum, can then be maximized to estimate the source position ϕ_0.

(c) Give the expression for the filter $\mathbf{w}$ related to the periodogram. Using (a), give the expression for $P(\phi)$. What is the maximum of this spectrum and for which value of ϕ is it obtained?

(d) Give the expression for the filter $\mathbf{w}$ related to the minimum variance spectral estimation method. Using (a), give the expression for $P(\phi)$. What is the maximum of this spectrum and for which value of ϕ is it obtained?

Hint: You will have to use a specific form of the matrix inversion lemma, i.e., $(c + \mathbf{x}\mathbf{x}^T)^{-1} = c^{-1}\mathbf{I} - c^{-1}(c + \mathbf{x}^T\mathbf{x})^{-1}\mathbf{x}\mathbf{x}^T$.

For finding the filter related to the MUSIC method, we solve the following problem. For every source position ϕ we minimize the distance between $\mathbf{w}$ and $\mathbf{a}(\phi)$ subject to the fact that $\mathbf{w}$ is orthogonal to the noise subspace of $\mathbf{R}_x$ described by the $K \times (K - r)$ matrix $\mathbf{U}_n$ [note that r is the rank you have to derive in (b)]. This problem can be mathematically formulated as

$$\min_{\mathbf{w}} \|\mathbf{w} - \mathbf{a}(\phi)\|^2 \quad \text{subject to} \quad \mathbf{U}_n^T\mathbf{w} = \mathbf{0}. \tag{2}$$

(e) Solve problem (2) using the method of Lagrange multipliers. Note that you have $K - r$ linear constraints so your Lagrange multiplier is a vector $\boldsymbol{\lambda}$ of size $K - r$. Prove that the solution is given by $\mathbf{w} = (\mathbf{I} - \mathbf{U}_n\mathbf{U}_n^T)\mathbf{a}(\phi)$.

(f) The useful signal energy of this MUSIC filter for a source at position ϕ is given by $|\mathbf{w}^T\mathbf{a}(\phi)|^2$. Show that maximizing this useful signal energy is the same as maximizing the so-called MUSIC spectrum given by $[\mathbf{a}^T(\phi)\mathbf{U}_n\mathbf{U}_n^T\mathbf{a}(\phi)]^{-1}$.