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

EE4C03 STATISTICAL DIGITAL SIGNAL PROCESSING AND MODELING

18 January 2021, 9:00–12:20

Block 1 (9:00-10:30)

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 10:25–10:35

This block consists of three questions (24 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 $x(n) = A\sin(\omega_0 n + \phi)$, where A and ω_0 are fixed constants, and ϕ is a random phase, uniformly distributed over $[-\pi, \pi]$. 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 autocovariance sequence $c_y(k)$.

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

$$r(k) = b\,a^{|k|}\,, \qquad (|a| < 1)$$

and construct the $N \times N$ correlation matrix $\mathbf{R}$. Take $N > 2$.

– Derive the rank of $\mathbf{R}$.

– How can a random signal with this autocorrelation sequence be generated from white noise?

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

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

(Motivate your answer.)

(d) *True* or *False:* if $y(n) = x_1(n) + x_2(n)$, and $x_1(n)$ is a WSS process and $x_2(n)$ is a WSS process and not independent of $x_1(n)$, then $y(n)$ is a WSS process. (Motivate your answer: give a proof or give a counterexample.)

Solution

(a)

$$c_x(k) = \frac{1}{2}A^2\cos(\omega_0 k) \qquad \Leftrightarrow \qquad P_x(\omega) = \frac{1}{2}\pi A^2\left[u_0(\omega-\omega_0) + u_0(\omega+\omega_0)\right]$$

where $u_0(\omega)$ is a delta spike in frequency domain.

$$h(k) = \delta(n) - \frac{1}{2}\delta(n-1) \quad \Leftrightarrow \quad H(z) = 1 - \frac{1}{2}z^{-1} \quad \Leftrightarrow \quad |H(\omega)|^2 = (1-\frac{1}{2}e^{-j\omega})(1-\frac{1}{2}e^{j\omega}) = \frac{5}{4} - \cos(\omega)$$

$$\begin{aligned} P_y(\omega) &= (\frac{5}{4} - \cos(\omega))\frac{1}{2}\pi A^2\left[u_0(\omega-\omega_0) + u_0(\omega+\omega_0)\right] \\ &= \frac{1}{2}\pi A^2\left\{\frac{5}{4} - \cos(\omega_0)\right\}\left[u_0(\omega-\omega_0) + u_0(\omega+\omega_0)\right] \end{aligned}$$

Hence

$$c_y(k) = \frac{1}{2}A^2\left\{\frac{5}{4} - \cos(\omega_0)\right\}\cos(\omega_0 k)$$

(b) The covariance matrix is

$$\mathbf{R} = b\begin{bmatrix} 1 & a & a^2 & \cdots & a^{N-1} \\ a & 1 & a & \cdots & a^{N-2} \\ a^2 & a & 1 & \cdots & a^{N-3} \\ \vdots & \vdots & & & \vdots \\ a^{N-1} & a^{N-2} & \cdots & a & 1 \end{bmatrix}$$

Since $|a| < 1$, the rank of $\mathbf{R}$ is N (full rank).

A signal with this autocorrelation can be generated by filtering white Gaussian noise with a first-order filter

$$H(z) = \frac{B}{1 - Az^{-1}}$$

Comparing to eqn (3.122), the resulting autocorrelation sequence is

$$r_x(k) = \frac{B}{1-A^2}A^{|k|}$$

we find that $A = a$, and $B = \sqrt{b(1-a^2)}$.

(c) *True.* The matrix is Hermitian, Toeplitz, and $r(0) \geq |r(k)|$. Because $r(0) = 2$, $r(2) = 2$, the process is periodic with period 2. The given values in $\mathbf{R}$ are consistent with this.

With Matlab, you could also compute the eigenvalues of $\mathbf{R}$ and note that they are all semi-positive. Actually, the rank is 2 (columns are repeated), and the covariance sequence is $r(k) = 1 + \cos(k\pi)$, so this could have been generated by a non-zero mean sinusoidal process.

(d) *False,* in general. The two processes $x_1(n)$ and $x_2(n)$ need to be *jointly* WSS, and this was not specified. In particular, $r_{x_1,x_2}(k,l)$ should depend only on the difference $k - l$.

A counterexample would be a case where $x_2(n) = x_1(-n)$.

Question 2 (9 points)

A random sequence $x(n)$ has the autocorrelation sequence

$$r_x(0) = 1,\ r_x(1) = 0.8,\ r_x(2) = 0.5,\ r_x(3) = 0.1$$

(a) Use the Schur recursion to compute, step by step, the reflection coefficients $\Gamma_1, \Gamma_2, \Gamma_3$ and the modeling errors $\epsilon_1, \epsilon_2, \epsilon_3$.

(b) Draw the corresponding lattice filter schematic that filters the signal $x(n)$ into white noise. (Clearly indicate the inputs and outputs.)

(c) From the reflection coefficients, compute the model parameters $a_3(k)$ of a 3rd order model.

(d) Write down the 3×3 Yule-Walker equations and verify your result under (c).

(e) If we would have $\Gamma_3 = 0$, what conclusions can you draw from this? (What can you say about the random process, what can you say about $\Gamma_4, \Gamma_5, \cdots$ and $\epsilon_4, \epsilon_5, \cdots$.)

Solution

(a)

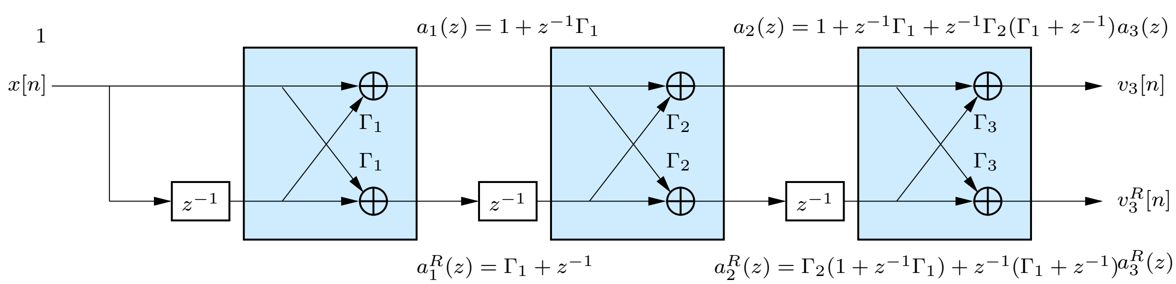
$$\begin{bmatrix} 0 & 1 \\ 0.8 & 0.8 \\ 0.5 & 0.5 \\ 0.1 & 0.1 \end{bmatrix} \underset{\text{shift}}{\Longrightarrow} \begin{bmatrix} 0 & 0 \\ 0.8 & 1 \\ 0.5 & 0.8 \\ 0.1 & 0.5 \end{bmatrix} \underset{\text{rotate, } \Gamma_1 = -0.8}{\Longrightarrow} \begin{bmatrix} 0 & 0 \\ 0 & 0.36 \\ -0.14 & 0.4 \\ -0.3 & 0.42 \end{bmatrix}$$

$$\underset{\text{shift}}{\Longrightarrow} \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ -0.14 & 0.36 \\ -0.3 & 0.4 \end{bmatrix} \underset{\text{rotate, } \Gamma_2 = \frac{0.14}{0.36} = 0.3889}{\Longrightarrow} \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0.3056 \\ -0.1444 & 0.2833 \end{bmatrix} \underset{\text{shift}}{\Longrightarrow} \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ -0.1444 & 0.3056 \end{bmatrix}$$

$$\underset{\text{rotate, } \Gamma_3 = \frac{0.1444}{0.3056} = 0.4725}{\Longrightarrow} \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0.2373 \end{bmatrix}$$

The modeling errors are $\epsilon_1 = 0.36,\ \epsilon_2 = 0.3056,\ \epsilon_3 = 0.2373$.

(b) See Levinson slides, p.18:



If the input is $x(n)$, the outputs are $v_3(n)$ (white noise up to order 3) and $v_3^R(n)$.

(c) If the input is $[1, 0, 0, 0]$ (a delta spike), the output at the top right is $a_3(n)$. Using the diagram, this can be computed by following the same procedure as above:

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \underset{\text{shift}}{\Longrightarrow} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \underset{\text{rotate, } \Gamma_1 = -0.8}{\Longrightarrow} \begin{bmatrix} 1 & -0.8 \\ -0.8 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\underset{\text{shift}}{\Longrightarrow} \begin{bmatrix} 1 & 0 \\ -0.8 & -0.8 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \underset{\text{rotate, } \Gamma_2 = \frac{0.14}{0.36} = 0.3889}{\Longrightarrow} \begin{bmatrix} 1 & 0.3889 \\ -1.1111 & -1.1111 \\ 0.3889 & 1 \\ 0 & 0 \end{bmatrix} \underset{\text{shift}}{\Longrightarrow} \begin{bmatrix} 1 & 0 \\ -1.1111 & 0.3889 \\ 0.3889 & -1.1111 \\ 0 & 1 \end{bmatrix}$$

$$\underset{\text{rotate, } \Gamma_3 = \frac{0.1444}{0.3056} = 0.4725}{\Longrightarrow} \begin{bmatrix} 1 & 0.4725 \\ -0.9274 & -0.1361 \\ -0.1361 & -0.9274 \\ 0.4725 & 1 \end{bmatrix}$$

The model parameters are $a_3(k) = [1, -0.9274, -0.1361, 0.4725]$.

(d) The Yule-Walker equations with the $a_3(k)$-coefficients filled in are

$$\begin{bmatrix} 1 & 0.8 & 0.5 \\ 0.8 & 1 & 0.8 \\ 0.5 & 0.8 & 1 \end{bmatrix} \begin{bmatrix} -0.9274 \\ -0.1361 \\ 0.4725 \end{bmatrix} = - \begin{bmatrix} 0.8 \\ 0.5 \\ 0.1 \end{bmatrix}$$

and indeed the LHS is equal to the RHS.

(e) If $\Gamma_3 = 0$, then the underlying process has an AR model of order 2. The prediction error $\epsilon_4 = \epsilon_3$, and becomes constant after the 2nd step. Γ_3 and all subsequent Γ_k will be zero: the recursion stops.

Question 3 (7 points)

We receive 1 second of a signal $x_a(t)$, sampled at a rate $f_s = 1000$ Hz to obtain a data sequence $x(n)$, $n = 0, \cdots, N-1$. The signal is assumed to be white Gaussian noise with variance σ^2 per sample, but it may be contaminated by a narrowband interferer. To detect it, Bartlett's method is used to estimate the power spectrum of $x(n)$. At each frequency we wish to obtain a variance of the estimate that is smaller than $\frac{1}{5}\sigma^4$.

(a) How should the number of segments K and the length of each segment L be chosen?

(b) What is best resolution (in Hz) that can be obtained using this technique?

(c) How can the resolution be improved? Give 3 possibilities. For each possibility, discuss the effect on the variance of the spectrum estimate.

(d) Suppose the bandwidth of the interferer is B Hz. To detect the interferer, does it help to improve the resolution? To answer this, consider the case where the resolution is smaller than B as well as the case where it is larger.

Solution

(a) For Bartlett's method, we have $N = KL$; the resolution is

$$\delta\omega = 0.89K\frac{2\pi}{N} = 0.89\frac{2\pi}{L}$$

and the variance is

$$\text{Var}(\hat{P}_B(e^{j\omega})) \approx \frac{1}{K}P_x^2(e^{j\omega})$$

In this case, we have $N = 1000$; the condition on the variance gives K: for those frequencies where there is no interferer, hence $P_x(e^{j\omega}) = \sigma^2$, follows $K = 5$.

From $N = 1000$ and $K = 5$, it follows that $L = 200$.

(b) The resolution is (for normalized radial frequencies ω)

$$\Delta\omega = 0.89K\frac{2\pi}{N} = 0.89\frac{2\pi}{L}$$

Translating to the original analog domain frequencies (F, in Hz, where $\omega = 2\pi f = 2\pi F/f_s$), we find

$$\Delta F = 0.89\frac{f_s}{L} = 4.45\,\text{Hz}$$

Note that we can also write $f_s = N/T$ where $T = 1$ sec is the observation time; it follows

$$\Delta F = 0.89\frac{N}{LT} = 0.89\frac{K}{T}$$

(c) 1. Increase L; with constant N we will have smaller K and hence a larger variance.

2. Decrease f_s; with a constant total measurement time we will have fewer samples N, if we keep L the same then K will decrease, hence the variance will be larger.

3. Increase the observation time; with the same f_s we will have a larger N, hence we can choose a larger L while K remains the same. With this option the variance can stay the same.

(d) The area under the peak in the spectral density estimate is the energy of the interferer.

1. Suppose the interferer bandwidth is smaller than the resolution. Then, if we halve the resolution (double L), the new frequency bin that contains the interferer will be half as wide. With the same interferer power, the 'peak' in the power spectrum will double, and we will more easily detect it. Thus, this seems useful. With constant N, the variance of the estimates will also double, visible in the plot by an increased standard deviation by a factor $\sqrt{2}$. So the effect will be less pronounced but still useful.

2. If the interferer bandwidth is larger than the resolution, then if we further improve the resolution, the peak will not increase anymore; it will not help to detect the interferer. If the improvement in resolution causes a larger variance of the estimates, the overall effect is even detrimental.

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

EE4C03 STATISTICAL DIGITAL SIGNAL PROCESSING AND MODELING

18 January 2021, 9:00–12:20

Block 2 (10:50-12:20)

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 12:15–12:25

This block consists of three questions (26 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 process $x(n)$ consisting of a sinusoid plus noise. Specifically, $x(n)$ is given by

$$x(n) = A(n)\cos(\omega n) + e(n),$$

where ω is a deterministic frequency, $A(n)$ is Gaussian with mean zero and variance σ_A^2, and $e(n)$ is white Gaussian with mean zero and variance σ_e^2. We assume $A(n)$ and $e(n)$ are independent. We want to estimate the amplitude $A(n)$ of this sinusoid by means of a simple zeroth order filter w applied to $x(n)$.

(a) What is the distribution of $x(n)$? What is its mean and its variance? Is $x(n)$ stationary?

(b) Whether $x(n)$ is stationary or not, can you think of a good filter w to estimate $A(n)$? In your answer, you may assume w is time-varying.

(c) Derive the normalized least mean squares (NLMS) algorithm to adaptively track the filter w that allows us to estimate $A(n)$ from $x(n)$.

(d) Suppose we now want to estimate $A(n)$ from $x(n)$ using a Kalman filter. In that case, we first need to make some assumption on the dynamics of $A(n)$. Since we have no specific information about the dynamics, we simply assume

$$A(n+1) = A(n).$$

Using the above equation as state equation, derive the Kalman filter to estimate $A(n)$ from $x(n)$. Name at least two advantages of this Kalman filter over the NLMS filter described in (c)?

Solution

(a) The process $x(n)$ is Gaussian with zero mean and variance $\sigma_A^2 \cos^2(\omega n) + \sigma_e^2$. Since the variance depends on n the process is not stationary.

(b) Since the process is not stationary, strictly speaking there is no optimal Wiener filter. If we would have to conisder a time-invariant filter w, the best we can do is to take $w = 0$, which is not very useful. A better time-varying choice could be $w^{(n)} = 1/\cos(\omega n)$, or a Wiener-style filter like

$$w^{(n)} = \frac{E\{A(n)x(n)\}}{E\{|x(n)|^2\}} = \frac{\sigma_A^2 \cos(\omega n)}{\sigma_A^2 \cos^2(\omega n) + \sigma_e^2}.$$

(c) The NLMS filter update looks as follows:

$$\hat{w}^{(n+1)} = \hat{w}^{(n)} - \frac{\mu}{|x(n)|^2} x(n)(\hat{w}^{(n)} x(n) - A(n)).$$

(d) The Kalman filter is based on the state space model

$$\begin{aligned} A(n+1) &= A(n), \\ x(n) &= A(n)\cos(\omega n) + e(n). \end{aligned}$$

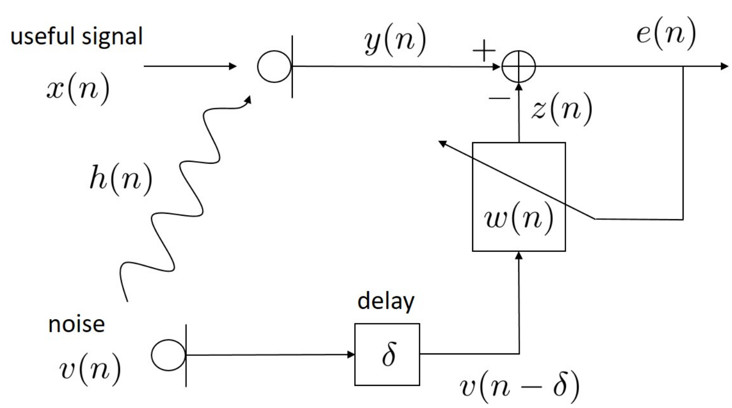
The updating equations are then given by

$$\begin{aligned} \hat{A}(n|n-1) &= \hat{A}(n-1|n-1) \\ p(n|n-1) &= p(n-1|n-1) \\ k(n) &= p(n|n-1)\cos(\omega n)[\cos^2(\omega n)p(n|n-1) + \sigma_e^2]^{-1} \\ \hat{A}(n|n) &= \hat{A}(n|n-1) + k(n)[x(n) - \hat{A}(n|n-1)\cos(\omega n)] \\ p(n|n) &= [1 - k(n)\cos(\omega n)]p(n|n-1) \end{aligned}$$

As far as the advantages is concerned, first of all, the Kalman filter provides the optimal estimate of the amplitude, given the prior data, whereas the NLMS filter does not. Second, it explicitly takes the time-varying behavior of the measurement model into account. Thirdly, it does not require the true reference signal $A(n)$ in the update equations.

Question 5 (8 points)

Consider the noise canceling scenario depicted below.



The useful signal $x(n)$ that is picked up by the microphone is corrupted by noise $v(n)$ after it passes through the noise path $h(n)$. So the signal $y(n)$ measured by the microphone is $y(n) = x(n) + h(n) * v(n)$, where $*$ represents the convolution operator. To cancel the noise, we apply a separate microphone at the noise source, which only picks up the noise $v(n)$. We let the noise $v(n)$ pass through the filter $w(n)$ and subtract its output $z(n)$ from $y(n)$ to obtain $e(n) = y(n) - z(n)$. Minimizing the error $e(n)$ we hope that the filter $w(n)$ will be close to the noise path $h(n)$ such that $e(n)$ will be close to the desired signal $x(n)$. We assume that the noise $v(n)$ is white Gaussian with zero mean and variance σ^2 and that it is independent from the useful signal $x(n)$.

Suppose we can only afford estimating a filter $w(n)$ of a short order p which is much smaller than the order of $h(n)$. Hence, the filter taps of $w(n)$ can be grouped in a short vector $\mathbf{w} = [w(0), w(1), \dots, w(p)]^T$. To compenstate for this limited order, we include a delay δ before the filter $w(n)$ and instead of filtering $v(n)$ we filter $v(n-\delta)$.

(a) First, express the optimal Wiener filter $\mathbf{w}$ as a function of $r_v(k) = E\{v(n)v(n-k)\}$ and $r_{yv}(k) = E\{y(n)v(n-k)\}$.

(b) Using the fact that $v(n)$ is white Gaussian with zero mean and variance σ^2, express $r_v(k)$ and $r_{yv}(k)$ as a function of $h(n)$.

(c) From the expressions in (b), write the optimal solution of $\mathbf{w}$ as a function of $h(n)$ and give the expression for the optimal error $e(n)$.

(d) Assume that we achieve our goal and $w(n) = h(n+\delta)$, for $n = 0, 1, \dots, p$. In that case, what would be a good way to select δ?

(e) Is this still a good approach when $v(n)$ is not white? Also explain why. Describe in words what will change.

Solution

(a) Defining the vector $\mathbf{v}_\delta(n) = [v(n-\delta), v(n-1-\delta), \dots, v(n-p-\delta)]^T$, the output of the filter $\mathbf{w}$ is given by $z(n) = \mathbf{w}^T\mathbf{v}_\delta(n)$. The optimal Wiener filter is this given by

$$\begin{aligned}\mathbf{w} &= E\{\mathbf{v}_\delta(n)\mathbf{v}_\delta^T(n)\}^{-1}E\{y(n)\mathbf{v}_\delta(n)\}\\ &= \begin{bmatrix} r_v(0) & \dots & r_v(p) \\ \vdots & \ddots \vdots & \\ r_v(-p) & \dots & r_v(0)\end{bmatrix}^{-1}\begin{bmatrix} r_{yv}(\delta) \\ \vdots \\ r_{yv}(p+\delta)\end{bmatrix}\end{aligned}$$

(b) Using the fact that $v(n)$ is white Gaussian noise with zero mean and variance σ^2, we have

$$\begin{aligned} r_v(k) &= \sigma^2\delta(k),\\ r_{yv}(k) &= E\{(x(n) + h(n) * v(n))v(n-k)\}\\ &= E\{(h(n) * v(n))v(n-k)\}\\ &= h(k) * r_v(k) = \sigma^2 h(k) * \delta(k) = \sigma^2 h(k).\end{aligned}$$

(c) From (b), it is easy to see that

$$\mathbf{w} = \begin{bmatrix} r_v(0) & \dots & r_v(p) \\ \vdots & \ddots & \vdots \\ r_v(-p) & \dots & r_v(0) \end{bmatrix}^{-1} \begin{bmatrix} r_{yv}(\delta) \\ \vdots \\ r_{yv}(p+\delta) \end{bmatrix}$$

$$= (\sigma^2 \mathbf{I})^{-1} \sigma^2 \begin{bmatrix} h(\delta) \\ \vdots \\ h(p+\delta) \end{bmatrix} = \begin{bmatrix} h(\delta) \\ \vdots \\ h(p+\delta) \end{bmatrix}.$$

The error is given by

$$e(n) = x(n) + h(n) * v(n) - \sum_{k=0}^{p} w(k) v(n-\delta-k)$$

$$= x(n) + h(n) * v(n) - \sum_{k=0}^{p} h(k+\delta) v(n-\delta-k).$$

(d) It seems logical to select a δ for which the mean square error is the smallest, i.e., for which $E\{e^2(n)\}$ is minimal. Since we can write

$$E\{e^2(n)\} = \text{constant} - \sigma^2 \sum_{k=0}^{p} w^2(k) = \text{constant} - \sigma^2 \sum_{k=0}^{p} h^2(k+\delta),$$

we select the δ for which $\|\mathbf{w}\|^2 = \sum_{k=0}^{p} w^2(k) = \sum_{k=0}^{p} h^2(k+\delta)$ is the highest. Selecting such a δ means the noise path $h(n)$ has maximal energy from the indices δ to $\delta + p$.

(e) Yes, the approach is then still valid. In that case, $w(n)$ will be approximately equal to $h(n+\delta)$ since also $r_v(k)$ will play a role.

Question 6 (9 points)

Consider a stochastic process $x(n)$ with zero mean and correlation function $r_x(k)$. Now assume we sample this process with two low-rate samplers, one at rate 1/3 and the other at rate 1/5. So the sequence obtained from the first sampler is $x_1(n) = x(3n)$ whereas the one related to the second sampler is $x_2(n) = x(5n)$.

(a) Express the correlation matrix of the first 5 samples of $x_1(n)$ as a function of $r_x(k)$, i.e., use $r_x(k)$ to compute $\mathbf{R}_{x_1} = E\{\mathbf{x}_1 \mathbf{x}_1^H\}$ with $\mathbf{x}_1 = [x(0), x(3), x(6), x(9), x(12)]^T$. Similarly express the correlation matrix of the first 3 samples of $x_2(n)$ as a function of $r_x(k)$, i.e., use $r_x(k)$ to compute $\mathbf{R}_{x_2} = E\{\mathbf{x}_2 \mathbf{x}_2^H\}$ with $\mathbf{x}_2 = [x(0), x(5), x(10)]^T$.

(b) Express the cross-correlation matrix between the first 5 samples of $x_1(n)$ and the 3 first samples of $x_2(n)$ as a function of $r_x(k)$, i.e., use $r_x(k)$ to compute $\mathbf{R}_{x_1,x_2} = E\{\mathbf{x}_1 \mathbf{x}_2^H\}$.

(c) How can these three correlation matrices be computed in practice if you have given $N_1 = 50$ samples of $x_1(n)$ and $N_2 = 30$ samples of $x_2(n)$?

(d) After computing the three correlation matrices $\mathbf{R}_{x_1}$, $\mathbf{R}_{x_2}$, and $\mathbf{R}_{x_1,x_2}$, which correlation values $r_x(k)$ can we estimate from $\mathbf{R}_{x_1}$, $\mathbf{R}_{x_2}$, and $\mathbf{R}_{x_1,x_2}$?

This result shows that even though we did not explicitly sample the process $x(n)$ at rate 1, we can still derive (part of) its correlation function from two low-rate samplers.

Now that we have estimated (part of) the correlation function $r_x(k)$, we can use this to interpolate the signals $x_1(n) = x(3n)$ and $x_2(n) = x(5n)$ in order to find the full rate 1 sequence $x(n)$. Assume for instance we want to estimate $x(n)$, for $n = 0, 1, 2, \ldots, 14$, from $\mathbf{x}_1$ and $\mathbf{x}_2$ jointly (or equivalently from $\mathbf{x}_0 = [\mathbf{x}_1^T, \mathbf{x}_2^T]^T$).

(e) Can you give the optimal Wiener filter expression to estimate $x(n)$, for $n = 0, 1, 2, \ldots, 14$, from $\mathbf{x}_1$ and $\mathbf{x}_2$? Try to express this filter as a function of $r_x(k)$.

(f) Do the correlation values from (d) provide enough information to estimate the full sequence $x(n)$, for $n = 0, 1, 2, \ldots, 14$, from $\mathbf{x}_1$ and $\mathbf{x}_2$? If so, explain why. If not, can you think of a way to circumvent this problem and still obtain a good estimate of $x(n)$, for $n = 0, 1, 2, \ldots, 14$?

Solution

(a) The correlation matrix of the first 5 samples of $x_1(n)$ is

$$\mathbf{R}_{x_1} = \begin{bmatrix} r_x(0) & r_x(3) & r_x(6) & r_x(9) & r_x(12) \\ r_x(-3) & r_x(0) & r_x(3) & r_x(6) & r_x(9) \\ r_x(-6) & r_x(-3) & r_x(0) & r_x(3) & r_x(6) \\ r_x(-9) & r_x(-6) & r_x(-3) & r_x(0) & r_x(3) \\ r_x(-12) & r_x(-9) & r_x(-6) & r_x(-3) & r_x(0) \end{bmatrix}$$

Similarly, the correlation matrix of the first 3 samples of $x_2(n)$ is

$$\mathbf{R}_{x_1} = \begin{bmatrix} r_x(0) & r_x(5) & r_x(10) \\ r_x(-5) & r_x(0) & r_x(5) \\ r_x(-10) & r_x(-5) & r_x(0) \end{bmatrix}$$

(b) The cross-correlation matrix between the first 5 samples of $x_1(n)$ and the 3 first samples of $x_2(n)$ is

$$\mathbf{R}_{x_1,x_2} = \begin{bmatrix} r_x(0) & r_x(5) & r_x(10) \\ r_x(-3) & r_x(2) & r_x(7) \\ r_x(-6) & r_x(-1) & r_x(4) \\ r_x(-9) & r_x(-4) & r_x(1) \\ r_x(-12) & r_x(-7) & r_x(-2) \end{bmatrix}$$

(c) In practice, the matrices can be computed as a temporal average. For the auto-correlation matrices, this is fairly easy. Defining $\mathbf{x}_1(n) = [x(3n), x(3n+3), x(3n+6), x(3n+9), x(3n+12)]^T$ and $\mathbf{x}_2(n) = [x(5n), x(5n+5), x(5n+10)]^T$, we can estimate them as

$$\hat{\mathbf{R}}_{x_1} = \frac{1}{50} \sum_{n=0}^{49} \mathbf{x}_1(n)\mathbf{x}_1^H(n),$$

$$\hat{\mathbf{R}}_{x_2} = \frac{1}{30} \sum_{n=0}^{29} \mathbf{x}_2(n)\mathbf{x}_2^H(n).$$

The cross-correlation matrix is a bit more tricky to compute since you have to make sure that the first samples of every block are lined up and this only happens every 15 samples.

So in this case we have to compute the average using the vectors $\tilde{\mathbf{x}}_1(n) = [x(15n), x(15n+3), x(15n+6), x(15n+9), x(15n+12)]^T$ and $\tilde{\mathbf{x}}_2(n) = [x(15n), x(15n+5), x(15n+10)]^T$, leading to

$$\hat{\mathbf{R}}_{x1,x2} = \frac{1}{10}\sum_{n=0}^{9} \tilde{\mathbf{x}}_1(n)\tilde{\mathbf{x}}_2^H(n).$$

Note that in this case we have less terms in the average.

(d) First note that if you have $r_x(-k)$, then you also know $r_x(k)$ since $r_x(-k) = r_x^*(k)$. As such, we know $r_x(k)$ for the lags $|k| = 0, 1, 2, 3, 4, 5, 6, 7, 9, 10, 12$.

(e) We look for the filter $\mathbf{w}_n$ to estimate $x(n)$ from $\mathbf{x}_0$, i.e., $\hat{\mathbf{x}}(n) = \mathbf{w}_n^T \mathbf{x}_0$. The Wiener filter is given by

$$\begin{aligned}\mathbf{w}_n &= E\{\mathbf{x}_0\mathbf{x}_0^H\}^{-1}E\{x(n)\mathbf{x}_0^*\} \\ &= \left[\begin{array}{c|c} \mathbf{R}_{x1} & \mathbf{R}_{x1,x2} \\ \hline \mathbf{R}_{x1,x2}^H & \mathbf{R}_{x2} \end{array}\right]^{-1} \left[\begin{array}{c} r_x(n) \\ r_x(n-3) \\ r_x(n-6) \\ r_x(n-9) \\ r_x(n-12) \\ \hline r_x(n) \\ r_x(n-5) \\ r_x(n-10) \end{array}\right]\end{aligned}$$

(f) We always know the correlation values inside the correlation matrix that is inverted. So that is not an issue. The question is whether we have available all correlation values inside $E\{x(n)\mathbf{x}_0^*\}$. Looking at these correlation values, it is clear that we would need all correlation values $r_x(k)$ for $|k| = 0, 1, 2, \ldots, 14$, in order to accurately estimate all samples $x(n)$, for $n = 0, 1, 2, \ldots, 14$. So what we have from (d) is not enough. To solve this problem, we could apply a Wiener filter on a part of $\mathbf{x}_0$ and not on the full vector. For instance to estimate $x(1)$, we need $r_x(1), r_x(-2), r_x(-5), r_x(-8), r_x(-11), r_x(-4), r_x(-9)$ if we want to apply the full blown filter. Since we don't have $r_x(-8)$ and $r_x(-11)$, we could drop the last two samples in $\mathbf{x}_1$ and only apply a Wiener filter on the other samples. That Wiener filter can be accurately computed. The same reasoning can be used for other $x(n)$ values.