**EE 779 Advanced Topics in Signal Processing**
**Assignment 1**
**Assigned: 07/08/13, Due: 20/08/13**
**Indian Institute of Technology Bombay**

## Note

- Please maintain academic integrity. Problems to be submitted should be your own work. Any copying will be penalized.
- Submit the starred $(*)$ problems and **all** simulations.
- For the simulation problems, you can use the Matlab functions provided by S & M [1] or Hayes [2]. A copy of the files is in moodle.

## Problems

1. Random process $v_1[n]$ and $v_2[n]$ are independent and have the same correlation function

$$r_v[n_1, n_0] = 0.5\delta[n_1 - n_0].$$

   (a) What is the correlation function of the random process

$$x[n] = v_1[n] + 2v_1[n+1] + 3v_2[n-1]?$$

   Is this random process WSS ?

   (b) Find the correlation matrix for a random vector consisting of eight consecutive samples of $x[n]$.

2. [$*$]A real random process has the correlation function

$$\mathbf{R}_x[l] = 2^{-|l|} + \delta[l].$$

   (a) What is the power spectral density (PSD) of this random process ?

   (b) Assume that the correlation function was estimated perfectly. A periodogram spectral estimate is formed by using only the first three lag values, that is,

$$\hat{r}_x[l] \begin{cases} r_x[l], & l = 0, \pm 1, \pm 2 \\ 0, & \text{otherwise.} \end{cases}$$

   What is the expression for the periodogram estimate ? Write it in simplest form.

3. A continuous-time signal $x_a(t)$ is bandlimited to 5 kHz, i.e., has a spectrum $X_a(f)$ that is zero for $|f| > 5$ kHz. Only 10 seconds of the signal has been recorded and is available for processing. We would like to estimate the power spectrum of $x_a(t)$ using the available data in a radix-2 FFT algorithm, and it is required that the estimate have a resolution of at least 10 Hz. Suppose that we use Bartlett's method of periodogram averaging.

   (a) If the data is sampled at Nyquist rate, what is the minimum section length that you may use to get the desired resolution ?

   (b) Using the minimum section length determined in part (a), with 10 seconds of data, how many sections are available for averaging ?

   (c) How does your choice of the sampling rate affect the resolution and variance of your estimate ? Are there any benefits to sampling above the Nyquist rate ?

4. Whenever the signal mean is unknown, a natural modification of the unbiased estimator of the autocorrelation function (ACF) is,

$$\tilde{r}_x(k) = \frac{1}{N-k} \sum_{n=k}^{N-1} \left(x(n) - \overline{x}\right)\left(x(n-k) - \overline{x}\right), \quad k = 0, \ldots, N-1. \tag{1}$$

and the biased estimator is,

$$\hat{r}_x(k) = \frac{1}{N} \sum_{n=k}^{N-1} \left(x(n) - \overline{x}\right)\left(x(n-k) - \overline{x}\right), \quad k = 0, \ldots, N-1. \tag{2}$$

where $\overline{x}$ is the sample mean

$$\overline{x} = \frac{1}{N} \sum_{n=0}^{N-1} x(n).$$

Show that in the unknown mean case, the usual names, "unbiased" and "biased" sample covariance associated with the above might no longer be appropriate. Indeed, both these estimators could be biased; furthermore, $\hat{r}_x(k)$ could be less biased than $\tilde{r}_x(k)$. To simplify calculations, assume that $x(n)$ is white noise.

5. [*]Show that the following definitions of the periodogram are equivalent.

$$P_{\text{per}}(e^{j\omega}) = \frac{1}{N} \left| \sum_{n=0}^{N-1} x(n)(e^{-j\omega n}) \right|^2. \tag{3}$$

$$P_{\text{per}}(e^{j\omega}) = \sum_{k=-(N-1)}^{N-1} \hat{r}_x(k) e^{-j\omega k}, \tag{4}$$

where $\hat{r}_x(k)$ is the biased autocorrelation estimate (assuming zero-mean $x(n)$).

6. Let $x(n), n = 0, \ldots, N-1$, be a real Gaussian process with zero mean and autocorrelation $r_x(k)$. The autocorrelation can be estimated using

$$\hat{r}_x(k) = \alpha(k) \sum_{n=k}^{N-1} \left(x(n) - \overline{x}\right)\left(x(n-k) - \overline{x}\right), \quad k = 0, \ldots, N-1, \tag{5}$$

where

$$\alpha(k) = \begin{cases} \frac{1}{N-k}, & \text{for unbiased ACF estimate} \\ \frac{1}{N}, & \text{for biased ACF estimate.} \end{cases}$$

Using the moment factoring theorem for real Guassian RVs (which is different from that for complex Gaussian RVs mentioned in class),

$$E\left\{abcd\right\} = E\left\{ab\right\} E\left\{cd\right\} + E\left\{ac\right\} E\left\{bd\right\} + E\left\{ad\right\} E\left\{bc\right\} - 2E\left\{a\right\} E\left\{b\right\} E\left\{c\right\} E\left\{d\right\}, \tag{6}$$

show that

$$Var\left\{\hat{r}_x(k)\right\} = \alpha^2(k) \sum_{m=-(N-k-1)}^{N-k-1} (N-k-|m|)\left(r_x^2(m) + r_x(m+k) r_x(m-k)\right).$$

**Simulations** For all simulation problems, you are advised to use Matlab's report generation facility and submit a html file. This should have proper comments and appropriately labelled plots. The analysis and observations are as important as writing the code. If you are using `scilab`, then you can submit a pdf file with appropriate comments and observation. A link will be provided in moodle to upload your final report/code.

1. Computer power spectrum estimates for the real data sets $S00$ and $S0X$ using the periodogram, Bartlett, Blackman-Tukey, and Welch procedures. In all cases zero-pad your data appropriately before giving it to the FFT routine so that you have computed enough points of the spectrum to give a smooth plot. Plotting 256 points of the spectrum should be sufficient.

   (a) For periodogram use N=16, 32, 128, and 512

   (b) For Bartlett, number of subsequences K = 32, 16, 4 (or size of subsequences L = 16, 32, 128)

   (c) For Welch, L = 16, 32, 128 and overlap L/2 (rectangular window)

   (d) For Blackman-Tukey, M = 16, 32, 128 (rectangular, Bartlett windows)

   Discuss the results of these experiments, especially trade-off between resolution and variance in the spectral estimate. Which choice do you think provides the best estimate of the true spectrum ? What can you say about the underlying processes ?

   Note: For the data set $S0X$, use $S01$ if your month of birth is Jan-Apr, $S02$ if your month of birth is May-Aug, and $S03$ if its Sep-Dec. Use the file `getfile.m` provided in moodle to read the data files.

## References

1. Petre Stoica and Randolph Moses, "Spectral analysis of signals", Prentice Hall, 2005. (Indian edition available)

2. Monson H. Hayes, "Statistical signal processing and modeling", Wiley India Pvt. Ltd., 2002. (Indian edition available)

3. Charles W. Therrien, "Discrete random signals and statistical signal processing", Charles W. Therrien, 2004.