Statistical Signal Processing

Midterm Exam

Thursday, March 28, 2012

Read Me First!

Only a personal cheat sheet is allowed.
No class notes, no exercise text or exercise solutions.

Write solutions on separate sheets,
***i.e.* no more than one solution per paper sheet.**

Return your sheets ordered according to problem (solution) numbering.

Return the text of the exam.

Warmup exercises

This is a warm up problem .. do not spend too much time on it.

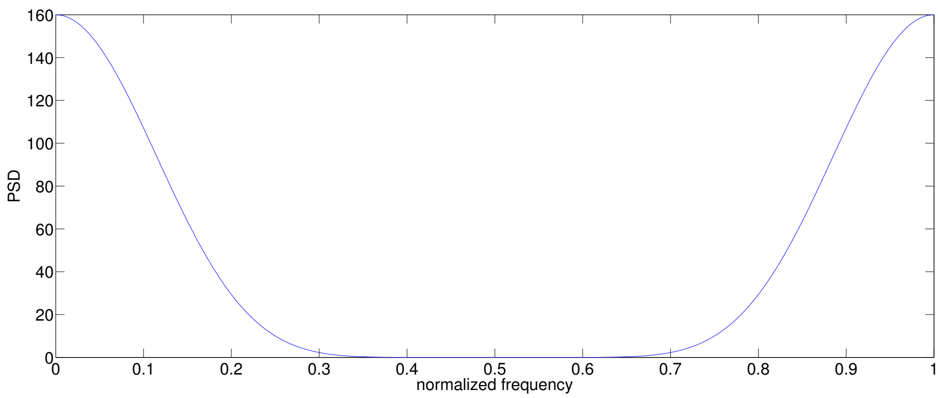
Please provide justified, rigorous, and simple answers.

Exercise 1. FILTERED WHITE NOISE (6 PTS)

A white noise $W[n]$ with variance $\sigma_W^2 = 1$ is filtered with a stable filter having transmittance $H(z)$.

1) Is the output a w.s.s. process? (justify precisely your answer)

2) Suppose that the output is a w.s.s. process, and that the power spectral density is given by



Draw on the z-plane the poles and zeros of the filter $H(z)$ (justify precisely your answer).

Solution 1.

Question 1

A white noise $W[n]$ is an i.i.d. stochastic process. Therefore

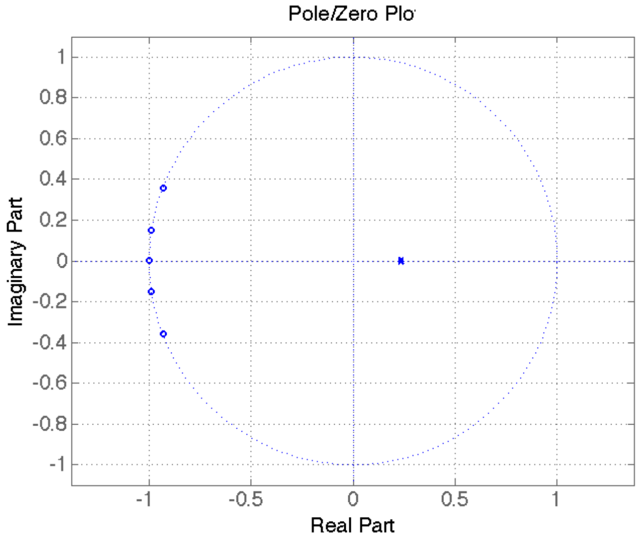
- $\mathbb{E}\left[W[n]\right] = \text{constant}$
- $\mathbb{E}\left[W[n+k]W[n]^*\right] = \delta_k \sigma_W^2$
- $\text{Var}\left(W[n]\right) = \sigma_W^2 = 1 < \infty$

Finally is a w.s.s. process.

Given that the input $W[n]$ of the filter is a w.s.s. process and that the filter $H(z)$ is stable, by the fundamental filtering formula the output $X[n]$ is also a w.s.s. process.

Question 2

An impulse response with square modulus given by the above plot corresponds to an impulse response with the following zeros and poles



That is

- Multiple zeros between $\pi/2$ and $-\pi/2$, to account for the fact that the impulse response is 0 for normalized frequencies between 0.35 and 0.65
- A pole on the real line between 0 and 1 (but strictly <1!)

Notice that you can use the Matlab tool *sptool* to generate filter and filter magnitude plots by placing poles and zeros on the z plane. Just be aware that Matlab normalizes the frequencies in [0,2] and be careful to correctly set the magnitude amplitude to linear and square.

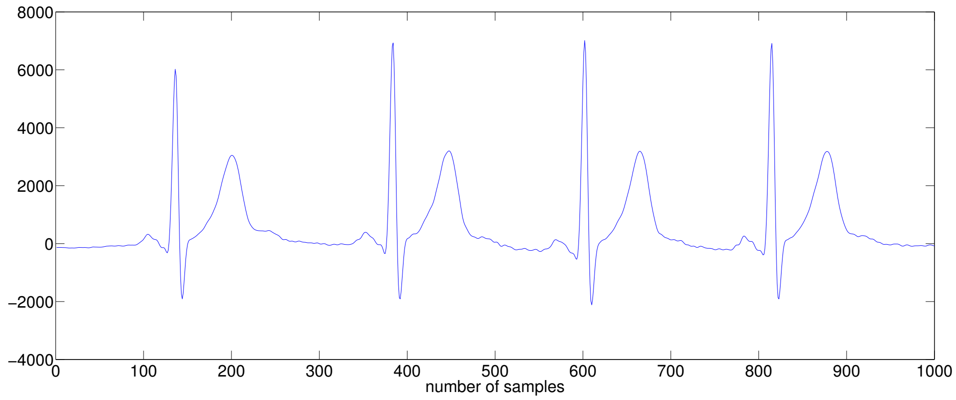
Main exercises

Here comes the core part of the exam .. take time to read the introduction and each problem statement.

Please provide justified, rigorous, and simple answers.

Exercise 2. Electrocardiogram - ECG (25 pts)

The electrocardiogram (ECG) is a signal representing the electrical activity of the heart.

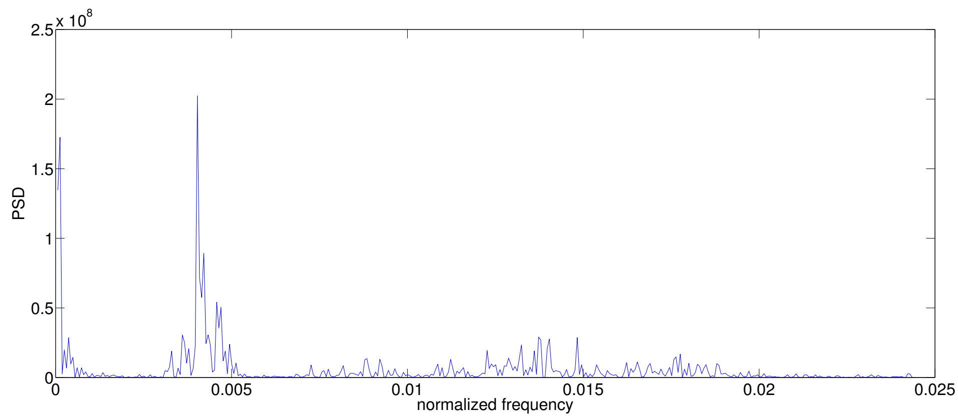


ECG signal.

It can be seen as a sequence of main pulses, where each pulse corresponds to a heart beat.

We model here the ECG as a stochastic process $X[n]$ that, over a short time interval (40 to 50 seconds), we assume to be wide sense stationary (w.s.s.).

The spectrum of an ECG provides relevant information on the heart activity and its analysis is, consequently, of foremost importance. A typical periodogram of an ECG signal is depicted in the figure below.



Periodogram of an ECG signal.

Recall that the number of pulses per minute is called the heart rate and that a normal heart rate of an adult goes from 40 to 200 beats per minute

1) Given that the sampling frequency is 250 Hz, based on the periodogram depicted above, can you give an estimate of the heart rate? (justify precisely your answer)

The periodogram depicted above has been computed using 10000 samples.

2) What is the spectral resolution in Hz? (justify precisely your answer)

Recall that, as discussed in class, the periodogram presents a high variance in the power spectral density estimation. Consequently, the obtained spectrum presents an irregular distribution of the energy. In addition, when analyzing an ECG spectrum, we are not interested in precisely detecting the heart rate but rather in evaluating the distribution of the energy around the frequency corresponding to the heart rate. Keeping this in mind, based on the information you can gather from the spectrum provided by the periodogram, you are asked to:

3) Propose a parametric spectral estimation method for the ECG signal (justify precisely your answer);
4) Describe in detail how to estimate the power spectral density with the parametric spectral estimation method you have proposed given that you have observed 10000 samples of the process $X[n]$.

In practice, due to the presence of electronic devices, an ECG signal $X[n]$ is often corrupted by noise. That is, the measured signal is actually

$$Y[n] = X[n] + W[n],$$

where $W[n]$ is a centered white Gaussian noise with variance σ_W^2.

Suppose to record 10000 samples (40 seconds) of the noisy ECG $Y[n]$,

5) Describe how we can de-noise such 10000 samples of $Y[n]$ using the Wiener filter (justify precisely your answer)

We now record the ECG signal for several hours.

6) Describe how we can de-noise several hours of recording of the ECG signal $Y[n]$ using the Wiener filter (justify precisely your answer)
hint: don' t forget that the Wiener filter works only on w.s.s. signals

Solution 2.

Question 1

A normal heart beat of 40 to 200 beats per minute corresponds to 0.6 to 3.3 beats per seconds. In Hz we have 0.6 to 3.3Hz and in normalized frequencies 0.0024 to 0.013. When we observe the power spectrum over the interval 0.0024 to 0.013 we see a maximum around the normalized frequency 0.004, which corresponds to 1Hz. Therefore we can estimate the heart rate to be of 60 beats per minute.

Question 2

The spectral resolution of the periodogram in normalized frequencies is given by $1/N$, where here $N = 10000$. Given that the sampling frequency is 250Hz, the spectral resolution is then given by$250/10000 = 0.025$ Hz.

Question 3

Here we are interested in the distribution of the energy at and around the frequency corresponding to the heart rate, and NOT in detecting the position and amplitude of the spectral peak with highest energy. Besides, as discussed in class, be careful that the peaky shape of the spectrum, and in particular the minor spectral lines appearing on the power spectrum, are mostly due to the error and the variance of the periodogram.

Therefore, we adopt a spectral estimation method for smooth spectrums, that is, a rational spectrum estimation based on Yule-Walker equations.

Question 4

We have observed N=10000 samples and we approximate the spectrum with a fractional polynomial of order $M \ll N$, that is $S_x(\omega) \simeq \left.\frac{1}{|P(z)|^2}\right|_{z=e^{j\omega}}$ where $P(z) = 1 + p_1 z^{-1} + \ldots + p_M z^{-M}$.

In order to estimate the polynomial coefficients, and therefore the spectrum, we need to

- Compute the empirical estimate of the correlation $\widehat{R}_X[k] = \frac{1}{N-k} \sum_{l=1}^{N-k} x[l+k]x[l]^*$, $k \geq 0$. For $k < 0$ we take $\widehat{R}_X[k] = \widehat{R}_X[-k]^*$. Notice that the fact of having $M << N$ reduces the error of the estimation of the correlation due to the extreme lags.
- Construct the system of Yule-Walker equations

$$\widehat{\boldsymbol{R}}_X^{M,M}(-\boldsymbol{p}^M) = \widehat{\boldsymbol{R}}_X^{1,M},$$

where

$$\widehat{\boldsymbol{R}}_X^{M,M} = \begin{bmatrix} \widehat{R}_X[0] & \widehat{R}_X[1] & \ldots & \widehat{R}_X[M-1] \\ \widehat{R}_X[1] & \widehat{R}_X[0] & \ldots & \widehat{R}_X[M-2] \\ \vdots & & \ddots & \vdots \\ \widehat{R}_X[M-1] & \widehat{R}_X[M-2] & \ldots & \widehat{R}_X[0] \end{bmatrix}$$

$$\boldsymbol{p}^M = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_M \end{bmatrix}, \quad \widehat{\boldsymbol{R}}_X^{1,M} = \begin{bmatrix} \widehat{R}_X[1] \\ \widehat{R}_X[2] \\ \vdots \\ \widehat{R}_X[M] \end{bmatrix}$$

- Solve the system (N^2 operations) to find $p_1, p_2, \ldots, p_M$.
- Obtain an estimate of the spectrum as

$$\widetilde{S}_x(\omega) = \left.\frac{1}{\left|\widetilde{P}(z)\right|^2}\right|_{z=e^{j\omega}} \quad \text{where } \widetilde{P}(z) = 1 + p_1 z^{-1} + \ldots + p_M z^{-M}$$

  obtained by solving the Yule-Walker system.

Question 5

We have $Y[n] = X[n] + W[n]$, where both $X[n]$ and $W[n]$ are w.s.s. processes over windows of 10000 samples. It follows that $Y[n]$ and $X[n]$ are jointly w.s.s. processes over windows of 10000 samples. Therefore, over windows of 10000 samples, we can apply the Wiener filter which is the standard de-noising tool in the presence of additive noise.

The Wiener filter $H(z)$ is given by $H(e^{jw}) = \dfrac{S_X(\omega)}{S_Y(\omega)}$, where $S_Y(\omega)$ can be estimated while $S_X(\omega)$ is unknown. Nevertheless, we have that $S_Y(\omega) = S_X(\omega) + \sigma_W^2$, which allows us to compute $S_X(\omega) = S_Y(\omega) - \sigma_W^2$. Finally, the Wiener Filter reads $H(e^{jw}) = \dfrac{S_Y(\omega) - \sigma_W^2}{S_Y(\omega)}$.

Question 6

Here we have to be careful since the Wiener filter works only on w.s.s. process and that the ECG signal can be assumed to be w.s.s. only over short time intervals of 40 to 50 seconds. Therefore we need to split the sever hours recording into segment of 40 seconds (10000 samples) and compute the Wiener filter for each segment. More precisely

- Decompose the several hours recording into segments of 10000 samples $\boldsymbol{Y}_k = [Y[1 + k10000], \ldots, Y[10000 + k10000]]$, $k = 0, 1, \ldots$.
- For each segment of 10000 $\boldsymbol{Y}_k$ compute the corresponding Wiener filter H_k (see Question 5). Here we will have a different filter for each segment!
- Apply the corresponding filter H_k so to de-noise each segment $\boldsymbol{Y}_k$ and obtain the de-noised segment $\widehat{\boldsymbol{X}}_k = [\widehat{X}[1 + k10000], \ldots, \widehat{X}[10000 + k10000]]$, $k = 0, 1, \ldots$
- Recompose the de-noised segments to form a de-noised version of the several hours ECG recording.

Exercise 3. ALTERNATE CODING (22 PTS)

Bits (0 and 1 values) are sent over a channel using an alternate signal, where +1 correspond to 1 and -1 to 0. We will consider two modeling approaches for the transmitted signal and the effect of the channel on it.

A) i.i.d. process and multiplicative noise

We start by modeling the transmitted signal as an i.i.d. stochastic process $X[n]$, taking values $\in \{-1, +1\}$ with equal probability.

1) Show that $X[n]$ is a w.s.s. process (justify precisely your answer).

When transmitted over the channel, the signal is corrupted by a multiplicative centered Gaussian noise $W[n]$ with unitary variance, *i.e.*, $\sigma_W^2 = 1$. The noise is assumed to be independent of the signal. Consequently the signal at the output of the channel is given by

$$Y[n] = W[n]X[n] \,.$$

2) Show that $Y[n]$ is a w.s.s. process (justify precisely your answer).
3) Compute the power spectral density (justify precisely your answer).
4) Give the distribution of the stochastic process $Y[n]$ (justify precisely your answer).
5) [bonus question] Give the joint distribution of $Y[n]$ and $X[n]$ (justify precisely your answer).

B) Markovian process and additive noise

We realize that modeling the transmission with an i.i.d. process corrupted by multiplicative noise is not really appropriate. We then switch to another model where we consider the signal $X[n]$ to be a Markov chain that is corrupted by an additive centered Gaussian noise $W[n]$ with unitary variance, *i.e.*, $\sigma_W^2 = 1$. Once again, the noise is assumed to be independent of the signal. Consequently the signal at the output of the channel is given by

$$Y[n] = X[n] + W[n] \,.$$

6) Give the distribution of $Y[1], \ldots, Y[N]$ (justify precisely your answer).
   hint: use Bayes' rule
7) How many parameters are necessary to characterize the process $Y[n]$? (justify precisely your answer).

Solution 3.

Question 1

The process is i.i.d., then

- $\mathbb{E}[X[n]] = \mathbb{E}[X[1]] = \mu = 0$ constant.
- $\mathbb{E}[X[n+k]X[n]^*] = \begin{cases} \mathbb{E}[|X[n]|^2] = \sigma_X^2 & \text{if } k = 0 \\ \mathbb{E}[X[n+k]]\,\mathbb{E}[X[n]^*] = 0 & \text{if } k \neq 0 \end{cases}.$

  Therefore $\mathbb{E}[X[n+k]X[n]^*] = \delta_k \sigma_X^2$ (depends only on the difference of the time lags).

Hence, $X[n]$ is a w.s.s. process

Question 2

We have

- $\mathbb{E}[Y[n]] = \mathbb{E}[X[n]W[n]] \overset{\text{independence}}{=} \mathbb{E}[X[n]]\,\mathbb{E}[W[n]] = 0$ constant.
- $\mathbb{E}[Y[n+k]Y[n]^*] = \mathbb{E}[X[n+k]W[n+k]X[n]^*W[n]^*] \overset{\text{independence}}{=}$ $\mathbb{E}[X[n+k]X[n]^*]\,\mathbb{E}[W[n+k]W[n]^*]$.

  Both $X[n]$ and $W[n]$ are i.i.d. processes so from Question 1 we have $\mathbb{E}[X[n+k]X[n]^*] = \delta_k \sigma_X^2$ and $\mathbb{E}[W[n+k]W[n]^*] = \delta_k \sigma_W^2$. Finally $\mathbb{E}[Y[n+k]Y[n]^*] = \delta_k \sigma_X^2 \sigma_W^2$.

Hence, $Y[n]$ is a w.s.s. process.

Question 3

The PSD $S_Y(\omega)$ is given by the Fourier transform of the correlation that here reads $R_Y[k] = \delta_k \sigma_X^2 \sigma_W^2$. Therefore $S_Y(\omega) = \sigma_X^2 \sigma_W^2$.

Question 4

The stochastic process $Y[n]$ is the product of two independent i.i.d. stochastic processes. Therefore is itself an i.i.d. stochastic process and to characterize its distribution it is sufficient to characterize the distribution of $Y[n]$ for a fixed n. For the sake of simplicity and without loss of generality we shall fix $n = 1$ and refer to $Y[1] = X[1]W[1]$ in the following.

Since $X[1]$ is a discrete value stochastic process and $W[1]$ is a continuous value stochastic process we shall use the cumulative distribution function $F_Y(a) = \mathbb{P}(Y[1] \leq a)$. Using the law to total probability and Baye's rule we have

$$\begin{aligned}
\mathbb{P}(Y[1] \leq a) &\overset{\text{total prob.}}{=} \sum_{k \in \{-1,1\}} \mathbb{P}(Y[1] \leq a, X[1] = k) \\
&\overset{\text{Baye's}}{=} \sum_{k \in \{-1,1\}} \mathbb{P}(Y[1] \leq a | X[1] = k)\,\mathbb{P}(X[1] = k) \\
&= \mathbb{P}(Y[1] \leq a | X[1] = -1)\,\mathbb{P}(X[1] = -1) + \mathbb{P}(Y[1] \leq a | X[1] = 1)\,\mathbb{P}(X[1] = 1) \\
&= \mathbb{P}(W[1] \leq -a)\,0.5 + \mathbb{P}(W[1] \leq a)\,0.5 = \mathbb{P}(W[1] \leq a)
\end{aligned}$$

where the latter equality is due to the symmetry of the centered Gaussian distribution ($\mathbb{P}(W[1] \leq -a) = \mathbb{P}(W[1] \leq a)$). So $Y[1]$ has a Gaussian distribution.

Question 6

Here $Y[n]$ is not an i.i.d. process. Therefore we need to consider the cumulative distribution of all the vector $Y[1], \ldots, Y[N]$, that is $F_{Y[1],\ldots,Y[N]}(a_1, \ldots, a_N) = \mathbb{P}(Y[1] \leq a_1, \ldots, Y[N] \leq a_N)$. Using the law to total probability and Baye's rule we have

$$\mathbb{P}(Y[1] \leq a_1, \ldots, Y[N] \leq a_N) \overset{\text{total prob.}}{=} \sum_{k_1,\ldots,k_N} \mathbb{P}(Y[1] \leq a_1, \ldots, Y[N] \leq a_N, X[1] = k_1, \ldots, X[N] = k_N) \overset{\text{Baye's}}{=}$$
$$\sum_{k_1,\ldots,k_N} \mathbb{P}(Y[1] \leq a_1, \ldots, Y[N] \leq a_N | X[1] = k_1, \ldots, X[N] = k_N) \mathbb{P}(X[1] = k_1, \ldots, X[N] = k_N)$$

Now

$$\begin{aligned}\mathbb{P}(Y[1] \leq a_1, \ldots, Y[N] \leq a_N | X[1] = k_1, \ldots, X[N] = k_N) \\ = \mathbb{P}(W[1] \leq a_1 - k_1, \ldots, W[N] \leq a_N - k_N) \\ \overset{\text{i.i.d.}}{=} \mathbb{P}(W[1] \leq a_1 - k_1) \mathbb{P}(W[2] \leq a_2 - k_2) \ldots \mathbb{P}(W[N] \leq a_N - k_N)\end{aligned}$$

and

$$\mathbb{P}(X[1] = k_1, \ldots, X[N] = k_N) \overset{\text{Baye's + Markov}}{=}$$
$$\mathbb{P}(X[N] = k_N | P[N-1] = k_{N-1}) \ldots \mathbb{P}(X[2] = k_2 | X[1] = k_1) \mathbb{P}(X[1] = k_1) = \pi_{k_1} p_{k_1,k_2} \cdots p_{k_{N-1},k_N}$$

So finally

$$\mathbb{P}(Y[1] \leq a_1, \ldots, Y[N] \leq a_N) = \sum_{k_1,\ldots,k_N} \mathbb{P}(W[1] \leq a_1 - k_1) \ldots \mathbb{P}(W[N] \leq a_N - k_N) \pi_{k_1} p_{k_1,k_2} \cdots p_{k_{N-1},k_N}$$

Question 7

Assuming the variance of the noise to be known (unitary as in the first art of the problem), the parameters that characterize the process $Y[n]$ are the parameters of the two states Markov chain, namely, the initial probabilities (π_{-1} and π_1) and the transition probabilities ($p_{1,1}$, $p_{1,-1}$, $p_{-1,1}$, and $p_{-1,-1}$). Taking into account the constraints

- $\pi_{-1} + \pi_1 = 1$
- $p_{1,1} + p_{1,-1} = 1$ and $p_{-1,1} + p_{-1,-1} = 1$

the process $Y[n]$ is characterized by 3 parameters.