UCSD ECE 153 | Handout #26
Prof. Young-Han Kim | Friday, December 5, 2008

**Solutions to Old Final Exam II**

1. *Coin with random bias (20 points).* You are given a coin but are not told what its bias (probability of heads) is. You are told instead that the bias is the outcome of a random variable $P \sim \text{Unif}[0,1]$. Assume $P$ does not change during the sequence of tosses.

   (a) What is the probability that the first three flips are heads?

   (b) What is the probability that the second flip is heads given that the first flip is heads?

   **Solution:**

   (a) We have
   $$\text{P}\{\text{first three flips are heads}|P = p\} = p^3.$$
   Hence, by the law of total probability,
   $$\text{P}\{\text{first three flips are heads}\} = \int_0^1 p^3 dp = 1/4.$$

   (b) Since
   $$\text{P}\{\text{the first flip is heads}\} = \int_0^1 p\, dp = 1/2$$
   and
   $$\text{P}\{\text{first two flips are heads}\} = \int_0^1 p^2 dp = 1/3,$$
   we have
   $$\text{P}\{\text{second flip heads}|\text{first flip heads}\} = 2/3.$$

2. *Estimation (20 points).* Let $X_1$ and $X_2$ be independent identically distributed random variables. Let $Y = X_1 + X_2$.

   (a) Find $E[X_1 - X_2|Y]$.

   (b) Find the minimum mean squared error estimate of $X_1$ given an observed value of $Y = X_1 + X_2$. (Hint: Consider $E[X_1 + X_2|X_1 + X_2]$.)

   **Solution:**

   (a) By symmetry, $E[X_1|Y] = E[X_2|Y]$. Hence, $E[X_1 - X_2|Y] = 0$.

(b) On one hand, we have $E[X_1+X_2|X_1+X_2] = X_1+X_2 = Y$. On the other hand, we have $E[X_1+X_2|X_1+X_2] = E[X_1|Y]+E[X_2|Y] = 2E[X_1|Y]$. Hence, the MMSE estimate of $X_1$ given $Y$ is

$$E[X_1|Y] = \frac{Y}{2}.$$

3. *Stationary process (20 points).* Consider the Gaussian autoregressive random process

$$X_{k+1} = \frac{1}{3}X_k + Z_k, \quad k = 0, 1, 2, \ldots,$$

where $Z_0, Z_1, Z_2, \ldots$ are i.i.d. $\sim N(0, 1)$.

(a) Find the distribution on $X_0$ that makes this a stationary stochastic process.

(b) What is the resulting autocorrelation $R_X(n)$?

**Solution:**

(a) Consider $X_0 \sim N(0, 9/8)$. It is easy to see that $X_n \sim N(0, 9/8)$, which implies that $X_n$ is stationary (why?).

(b) We have

$$\begin{aligned} X_n &= \frac{1}{3}X_{n-1} + Z_{n-1} \\ &= \left(\frac{1}{3}\right)^2 X_{n-2} + \frac{1}{3}Z_{n-2} + Z_{n-1} \\ &= \left(\frac{1}{3}\right)^n X_0 + \sum_{k=1}^{n} \left(\frac{1}{3}\right)^{n-k} Z_{k-1}. \end{aligned}$$

Hence,

$$R_X(n) = EX_0X_n = \left(\frac{1}{3}\right)^n$$

for $n \geq 0$, and in general

$$R_X(n) = EX_0X_n = \left(\frac{1}{3}\right)^{|n|}.$$

4. *Prediction error process (20 points).* Let $X_1, X_2, \ldots$ be a discrete-time random process. Let $\hat{X}_n$ denote the linear MMSE estimator of $X_n$ given $X_1, X_2, \ldots, X_{n-1}$. Thus $\hat{X}_n$ is of the form $\hat{X}_n = \sum_{i=1}^{n-1} a_{ni}X_i + b_n$. Let $Z_n = X_n - \hat{X}_n$ denote the error.

(a) Find $EZ_n$.

(b) Find $\text{Cov}(Z_n, Z_m)$ for $m > n$. Interpret.

**Solution:**

(a) Obviously $EZ_n = 0$; otherwise we can reduce the mean square error further. (Why? Recall that $EZ_n^2 \geq E(Z_n - EZ_n)^2$.)

Alternatively, we see that the optimal choice of $b_n$ is given by

$$b_n = EX_n - \sum_{i=1}^{n-1} a_{ni} EX_i,$$

which implies that $EZ_n = 0$.

(b) Since $EZ_n = 0$, we simply consider $EZ_nZ_m$. Now by orthogonality principle, $Z_m$ is orthogonal to $(X_1, X_2, \ldots, X_{m-1})$. But $Z_n$ is a linear function of $(X_1, X_2, \ldots, X_n)$. Hence, $Z_m$ is orthogonal to $Z_n$, i.e., $EZ_nZ_m = 0$.

This shows that successive prediction errors are uncorrelated, which leads to a systematic method of representing a random process $X_n$ as

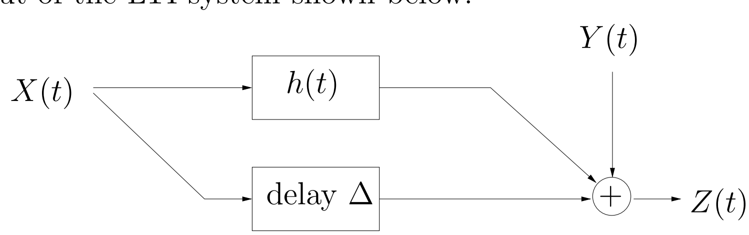
$$X_n = X_1 + \sum_{i=2}^{n} Z_i,$$

where the prediction error $Z_n$ can be interpreted as fresh randomness (innovation) at each time $n$. This representation is the basis for the so-called innovations approach to statistical signal processing.

5. *Minimum power (30 points).* Let $X(t)$ and $Y(t)$ be jointly WSS processes with zero mean. Let

$$\begin{aligned} Z(t) &= h(t) * X(t) + X(t - \Delta) + Y(t) \\ &= \int_{-\infty}^{\infty} h(s)X(t-s)ds + X(t-\Delta) + Y(t) \end{aligned}$$

be the output of the LTI system shown below.




(a) Find the power spectral density $S_Z(f)$ of the output $Z(t)$. Your answer should be in terms of $S_X(f)$, $S_Y(f)$, $S_{XY}(f)$, $\Delta$, and $H(f)$.

(b) Now find the LTI filter $H(f)$ that minimizes the power of $Z(t)$. (Hint: Consider $U(t) = -X(t-\Delta) - Y(t)$. Which filter minimizes the power of $Z(t) = h(t) * X(t) - U(t)$?)

(c) What is the resulting power $EZ^2(t)$?

**Solution:**

(a) We can write Z(t) as follows. $Z(t) = h(t)*X(t)+\delta(t-\Delta)*X(t)+Y(t) = (h(t)+\delta(t-\Delta))*X(t)+Y(t)$. Let $g(t) = h(t)+\delta(t-\Delta)$. Then, $Z(t) = (g(t)*X(t))+Y(t)$. To find $S_Z(f)$, we first find the autocorrelation function as

$$\begin{aligned} R_Z(\tau) &= E(Z(t)Z(t+\tau)) \\ &= E[(g(t) * X(t) + Y(t))(g(t+\tau) * X(t+\tau) + Y(t+\tau))] \\ &= R_{g*X}(\tau) + R_{g*X,Y}(\tau) + R_{Y,g*X}(\tau) + R_Y(\tau) \\ &= R_{g*X}(\tau) + R_{g*X,Y}(\tau) + R_{g*X,Y}(-\tau) + R_Y(\tau). \end{aligned}$$

But we have

$$\begin{aligned} R_{g*X,Y}(\tau) &= E[(g(t+\tau) * X(t+\tau))Y(t)] \\ &= E\left[Y(t)\int_{-\infty}^{\infty} g(r)X(t+\tau-r)dr\right] \\ &= \int_{-\infty}^{\infty} E[Y(t)X(t+\tau-r)]g(r)dr \\ &= \int_{-\infty}^{\infty} R_{XY}(\tau-r)g(r)dr = g(\tau) * R_{XY}(\tau). \end{aligned}$$

Thus,

$$\begin{aligned} S_Z(f) &= |G(f)|^2 S_X(f) + G(f)S_{XY}(f) + G^*(f)S^*_{XY}(f) + S_Y(f) \\ &= |G(f)|^2 S_X(f) + 2\mathrm{Re}[G(f)S_{XY}(f)] + S_Y(f) \\ &= |H(f) + e^{-j2\pi\Delta f}|^2 S_X(f) + 2\mathrm{Re}[(H(f) + e^{-j2\pi\Delta f})S_{XY}(f)] + S_Y(f). \end{aligned}$$

(b) Since $Z(t) = h(t)*X(t)-U(t)$, the LTI filter $h(t)$ that minimizes $EZ^2(t)$ is actually the optimum filter for estimating $U(t)$ based on the observation $\{X(s)\}_{-\infty<s<\infty}$ minimizing the MSE $E[(h(t) * X(t) - U(t))^2]$.
Therefore,

$$H(f) = \frac{S_{UX}(f)}{S_X(f)}.$$

Now consider

$$\begin{aligned} R_{UX}(\tau) &= E[(-X(t-\Delta+\tau) - Y(t+\tau))X(t)] \\ &= -E[X(t-\Delta+\tau)X(t)] - E[Y(t+\tau)X(t)] \\ &= -R_X(\Delta-\tau) - R_{YX}(\tau) \\ &= -R_X(\tau-\Delta) - R_{YX}(\tau). \end{aligned}$$

Thus,

$$\begin{aligned} S_{UX}(f) &= -e^{-j2\pi\Delta f} S_X(f) - S_{YX}(f) \\ &= -e^{-j2\pi\Delta f} S_X(f) - S^*_{XY}(f). \end{aligned}$$

By combining the above, we have

$$H(f) = \frac{S_{UX}(f)}{S_X(f)} = \frac{-e^{-j2\pi\Delta f} S_X(f) - S^*_{XY}(f)}{S_X(f)} = -e^{-j2\pi\Delta f} - \frac{S^*_{XY}(f)}{S_X(f)}.$$

Alternatively, we can optimize $G(f)$ instead of $H(f)$.

$$E[Z^2(t)] = E[(Y(t) + (h(t) + \delta(t-\Delta)) * X(t))^2] \quad = E[(Y(t) + g(t) * X(t))^2],$$

so that

$$-G(f) = \frac{S_{YX}(f)}{S_X(f)}.$$

Thus,

$$H(f) = G(f) - e^{-j2\pi\Delta f} = -e^{-j2\pi\Delta f} - \frac{S_{YX}(f)}{S_X(f)}.$$

(c) From the relationship between $R_Z(\tau)$ and $S_Z(f)$, we have

$$E[Z^2(t)] = R_Z(0) = \int_{-\infty}^{\infty} S_Z(f) df.$$

From parts (a) and (b),

$$\begin{aligned} S_Z(f) =& |H(f) + e^{-j2\pi\Delta f}|^2 S_X(f) + 2\mathrm{Re}[(H(f) + e^{-j2\pi\Delta f}) S_{XY}(f)] + S_Y(f) \\ =& \left| \frac{-e^{-j2\pi\Delta f} S_X(f) - S^*_{XY}(f)}{S_X(f)} + e^{-j2\pi\Delta f} \right|^2 S_X(f) \\ &+ 2\mathrm{Re}\left[ \left( \frac{-e^{-j2\pi\Delta f} S_X(f) - S^*_{XY}(f)}{S_X(f)} + e^{-j2\pi\Delta f} \right) S_{XY}(f) \right] + S_Y(f) \\ =& \left| \frac{-S^*_{XY}(f)}{S_X(f)} \right|^2 S_X(f) - 2\frac{|S_{XY}(f)|^2}{S_X(f)} + S_Y(f) \\ =& -\frac{|S_{XY}(f)|^2}{S_X(f)} + S_Y(f). \end{aligned}$$

Therefore,

$$\begin{aligned} E[Z^2(t)] &= \int_{-\infty}^{\infty} -\frac{|S_{XY}(f)|^2}{S_X(f)} + S_Y(f) df \\ &= E[Y^2(t)] - \int_{-\infty}^{\infty} \frac{|S_{XY}(f)|^2}{S_X(f)} df. \end{aligned}$$

This can be easily seen from the alternative approach in part (b), since the minimum power is the minimum MSE for estimating $Y(t)$ from $\{X(s)\}_{-\infty<s<\infty}$.

6. *Integration (40 points).* Let $X(t)$ be a zero-mean WSS process with autocorrelation function $R_X(\tau) = e^{-|\tau|}$. Let

$$Y(t) = \int_t^{t+1} X(s)ds.$$

(a) Is $Y(t)$ WSS?

(b) Is $(X(t), Y(t))$ jointly WSS?

(c) Find the linear MMSE estimate of $Y(t)$ given $X(t)$. Leave your answer in terms of $e$.

(d) Find the linear MMSE estimate of $Y(t)$ given $X(t)$ and $X(t+1)$.

**Solution:**

(a) Yes, $Y(t)$ is WSS. There are two ways to see this. Since the integration is a linear operation and

$$Y(t-t_0) = \int_{t-t_0}^{t-t_0+1} X(s)ds = \int_t^{t+1} X(s-t_0)ds,$$

we can regard $Y(t)$ as the output of an LTI system with WSS input $X(t)$. Therefore, $Y(t)$ is WSS.

Alternatively, we can check the WSS property of Y(t) by direct calculation.

$$E[Y(t)] = E\left[\int_t^{t+1} X(s)ds\right] = \int_t^{t+1} E[X(s)]ds = 0.$$

Consider

$$\begin{aligned}
R_Y(t_1, t_2) &= E[Y(t_1)Y(t_2)] \\
&= E\left[\int_{t_1}^{t_1+1} X(s)ds \int_{t_2}^{t_2+1} X(u)du\right] \\
&= \int_{t_1}^{t_1+1} \int_{t_2}^{t_2+1} E[X(s)X(u)]dsdu \\
&= \int_{t_1}^{t_1+1} \int_{t_2}^{t_2+1} R_X(s-u)dsdu \\
&= \int_0^1 \int_0^1 R_X(s-u+t_2-t_1)dsdu
\end{aligned}$$

Therefore, $R_Y(t_1, t_2)$ depends only on $t_2 - t_1$.

(b) Yes, $(X(t), Y(t))$ is jointly WSS. Again, since $Y(t)$ is the output of an LTI system with WSS input $X(t)$, they are jointly WSS. Alternatively, we can check the joint

WSS property of $(X(t), Y(t))$ by direct calculation. Consider

$$\begin{aligned} R_{XY}(t_1, t_2) &= E\left[X(t_1)\int_{t_2}^{t_2+1} X(s)ds\right] \\ &= \int_{t_2}^{t_2+1} E[X(t_1)X(s)]ds \\ &= \int_{t_2}^{t_2+1} R_X(s-t_1)ds \\ &= \int_0^1 R_X(s+t_2-t_1)ds, \end{aligned}$$

which depends only on $t_2 - t_1$. This proves that $(X(t), Y(t))$ is jointly WSS.

(c) We can apply the usual linear MMSE estimator formula to this part. We have

$$\sigma^2_{X(t)} = E[X^2(t)] = R_X(0) = 1,$$

$$\begin{aligned} \mathrm{Cov}(Y(t), X(t)) &= E[Y(t)X(t)] - 0 \\ &= R_{XY}(\tau = 0) \\ &= \int_0^1 e^{-s} ds \\ &= 1 - e^{-1}. \end{aligned}$$

Therefore,

$$\hat{Y}(t) = \frac{\mathrm{Cov}(Y(t), X(t))}{\sigma^2_{X(t)}}(X(t) - EX(t)) + EY(t) = (1 - e^{-1})X(t).$$

(d) Here we estimate $Y(t)$ from a random vector observation $\begin{bmatrix} X(t) & X(t+1) \end{bmatrix}^T$. Since $E(Y(t)X(t+1)) = R_{XY}(\tau = 1) = \int_0^1 e^{s-1} ds = 1 - e^{-1}$,

$$\begin{aligned} \hat{Y}(t) &= \begin{bmatrix} E[Y(t)X(t)] & E[Y(t)X(t+1)] \end{bmatrix} \begin{bmatrix} R_X(0) & R_X(1) \\ R_X(1) & R_X(0) \end{bmatrix}^{-1} \begin{bmatrix} X(t) \\ X(t+1) \end{bmatrix} \\ &= \begin{bmatrix} 1-e^{-1} & 1-e^{-1} \end{bmatrix} \begin{bmatrix} 1 & e^{-1} \\ e^{-1} & 1 \end{bmatrix}^{-1} \begin{bmatrix} X(t) \\ X(t+1) \end{bmatrix} \\ &= \frac{1-e^{-1}}{1+e^{-1}}(X(t) + X(t+1)). \end{aligned}$$