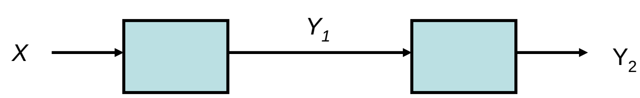
EECS 225A Spring 2005

Homework 4 solutions

---

1. As shown below, a random variable X is the input to a cascade of two systems with random variable outputs Y_1 and Y_2. You are given the joint PDF $p_{X,Y_1,Y_2}(x, y_1, y_2)$ and told that it satisfies $p_{Y_2|X,Y_1}(y_2 \mid x, y_1) = p_{Y_2|Y_1}(y_2 \mid y_1)$.
   a. Give an intuitive argument as to why this latter condition specifies that there are no 'hidden' information channels from X to Y_2, other than the path via Y_1.
   b. Suppose the MMSE optimum non-linear estimates of X based on Y_1 and Y_2 are, respectively, $f(y_1)$ and $g(y_2)$. In terms of the given PDF (or conditional PDF's derived from it) write down formulas for the functions $f(\cdot)$ and $g(\cdot)$.
   c. Show that $g(y_2) = E\left[f(Y_1) \mid Y_2 = y_2\right]$. Interpret this equation verbally, and give an intuitive argument for it.
   d. Show that $E\left|X - g(Y_2)\right|^2 = E\left|X - f(Y_1)\right|^2 + E\left|f(Y_1) - g(Y_2)\right|^2$. Interpret this equation verbally, and give an intuitive argument for it. Note that this relationship attributes the total MMSE to a contribution from each of the cascaded systems.
   e. Use the result of d. to attribute the total MMSE determined in Problem #4 of Homework #3 to the A/D converter and to the BSC. Use Matlab to plot the total and constituent MSE's as a function of crossover probability p for a Gaussian RV and interpret the result.



Solution

a. The condition specifies that conditional on knowledge of $Y_1 = y_1$, X and Y_2 are statistically independent. Thus, if we know Y_2 there is no additional statistical relationship to X that can be exploited.

b. $f(y_1) = \int x \cdot p_{X|Y_1}(x \mid y_1) \cdot dx$ and $g(y_2) = \int x \cdot p_{X|Y_2}(x \mid y_2) \cdot dx$

c. First, we can show that (pardon the shorthand notation):

$$p(x \mid y_1, y_2) = \frac{p(x, y_1, y_2)}{p(y_1, y_2)} = \frac{p(y_2 \mid x, y_1) \cdot p(x, y_1)}{p(y_2 \mid y_1) \cdot p(y_1)} = \frac{p(x, y_1)}{p(y_1)} = p(x \mid y_1)$$

Then we can rewrite $g(y_2)$ in terms of this density,

$$g(y_2) = \iint x \cdot p(x \mid y_1, y_2) p(y_1 \mid y_2) \cdot dy_1 \cdot dx = \iint x \cdot p(x \mid y_1) p(y_1 \mid y_2) \cdot dy_1 \cdot dx$$

or performing the x-integral,

$$g(y_2) = \int f(y_1) \cdot p(y_1 \mid y_2) \cdot dy_1 = E\left[f(Y_1) \mid Y_2 = y_2 \right]$$

The MMSE estimate of Y_2 can be formed from Y_1 alone, ignoring X. It is formed not by taking the conditional mean of Y_1, but rather the conditional mean of the MMSE estimator of X based on Y_1. Thus, that conditional mean is able to stand in as a proxy for X in calculating the new estimate.

d. We know from class that

$$E\left|X - f(Y_1)\right|^2 = E\left|X\right|^2 - E\left|f(Y_1)\right|^2$$

$$E\left|X - g(Y_2)\right|^2 = E\left|X\right|^2 - E\left|g(Y_2)\right|^2$$

Further, we can conclude that since $g(y_2) = E\left[f(Y_1) \mid Y_2 = y_2 \right]$ is the MMSE estimator of random variable $f(Y_1)$ based on observation of Y_2,

$$E\left| f(Y_1) - g(Y_2)\right|^2 = E\left| f(Y_1)\right|^2 - E\left|g(Y_2)\right|^2.$$

Combining these three equations leads to the desired result. This says if the intermediate RV Y_1 is replaced by the MMSE estimate of X based on Y_1, then the error in through the first and second systems are orthogonal and the MMSE of the cascaded systems is the sum of the MMSE's of the constituent systems. Without this replacement, the errors are in general correlated and thus we can't conclude anything in general about their individual contributions to the cascaded error.

e. Define the output of the A/D as Y_1 and the output of the BSC as Y_2. In class we showed that the MMSE estimator at the A/D output assumes values

$$f(Y_1) = \begin{cases} \mu_0, & Y_1 = 0 \\ \mu_1, & Y_1 = 1 \end{cases}$$

where

$$p_0 = \int_{-\infty}^{\alpha} p(x) \cdot dx, \ \mu_0 = \frac{1}{p_0} \cdot \int_{-\infty}^{\alpha} x \cdot p(x) \cdot dx \ \text{ and } \ \mu_1 = \frac{1}{(1 - p_0)} \cdot \int_{\alpha}^{\infty} x \cdot p(x) \cdot dx$$

Define a notation for the two values of the MMSE estimator at the BSC output,

$$g(Y_2) = \begin{cases} \tau_0, & Y_2 = 0 \\ \tau_1, & Y_2 = 1 \end{cases}.$$

These are the conditional means of $f(Y_1)$, conditioned on knowledge of Y_2. Calculating these conditional means,

$$q_0 = p_0(1 - p) + (1 - p_0)p$$

$$\tau_0 = E[f(Y_1) \mid Y_2 = 0] = \frac{\mu_0 p_0 (1 - p) + \mu_1 (1 - p_0) p}{q_0}$$

$$\tau_1 = E[f(Y_1) \mid Y_2 = 1] = \frac{\mu_0 p_0 p + \mu_1 (1-p_0)(1-p)}{1-q_0}$$

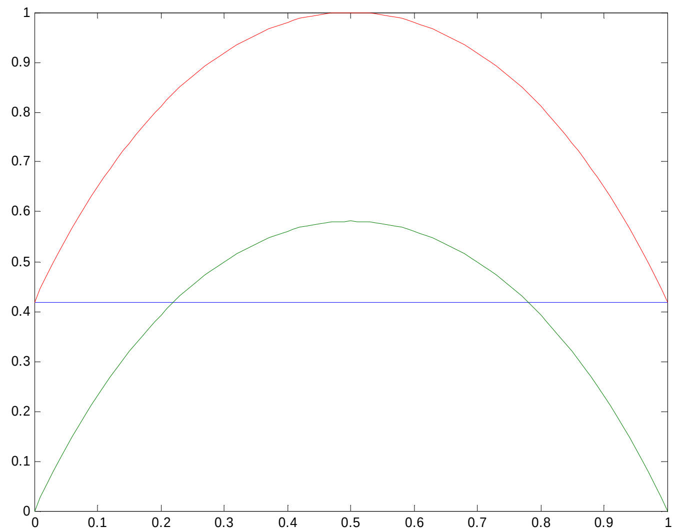
This is the same answer we calculated (using a more direct but cumbersome approach) in Homework #3.

To calculate the MMSE, all we need to know is the second moments,

$$E[f^2(Y_1)] = \mu_0^2 \cdot p_0 + \mu_1^2 \cdot (1-p_0)$$

$$E[g^2(Y_2)] = \tau_0^2 \cdot q_0 + \tau_1^2 \cdot (1-q_0)$$

See M-file hmwk04.m. The MMSE is plotted below for a zero-mean unit-variance Gaussian and threshold $\alpha = 0.5$. The blue line is the MMSE of the A/D, which is not dependent on p. The green line is the MMSE attributable to the BSC, which of course is zero at the end points and maximum at $p = 0.5$. The red line is the total MMSE, which goes to unity (the input variance) at $p = 0.5$ since in this case the output of the BSC is statistically independent of the input to the A/D.



---

2. Hayes Problem 3.3

Solution

(a) Since $x(n)$ is the output of an all-pole filter driven by white noise, $x(n)$ is an AR(p) process with a power spectrum

$$P_x(e^{j\omega}) = \frac{\sigma_w^2}{|A(e^{j\omega})|^2}$$

where

$$A(e^{j\omega}) = 1 - \sum_{k=1}^{p} a(k) e^{-jk\omega}$$

(b) The process $z(n)$ is a sum of two random processes

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

Since $x(n)$ is a linear combination of values of $w(n)$,

$$x(n) = \sum_{k=-\infty}^{n} h(k)w(n-k)$$

where $h(n)$ is the unit sample response of the filter generating $x(n)$, and since $v(n)$ is uncorrelated with $w(n)$, then $v(n)$ is uncorrelated with $x(n)$, and we have

$$r_z(k) = r_x(k) + r_v(k)$$

Therefore,

$$P_z(e^{j\omega}) = P_x(e^{j\omega}) + P_v(e^{j\omega})$$

and

$$P_z(e^{j\omega}) = \frac{\sigma_w^2}{|A(e^{j\omega})|^2} + \sigma_v^2 = \frac{\sigma_w^2 + \sigma_v^2|A(e^{j\omega})|^2}{|A(e^{j\omega})|^2}$$

---

3. Hayes Problem 3.4

Solution

(a) The power spectrum of $x(n)$ is

$$P_x(z) = \frac{3/4}{(1-\frac{1}{2}z^{-1})(1-\frac{1}{2}z)}$$

and the power spectrum of $y(n)$ is

$$P_y(z) = H(z)H(z^{-1})P_x(z) = \frac{3/4}{(1-\frac{1}{3}z^{-1})(1-\frac{1}{3}z)}$$

(b) The autocorrelation sequence for $y(n)$ may be easily found using the z-transform pair

$$\alpha^{|k|} \longleftrightarrow \frac{1-\alpha^2}{(1-\alpha z^{-1})(1-\alpha z)}$$

Since

$$(\tfrac{1}{3})^{|k|} \longleftrightarrow \frac{8/9}{(1-\frac{1}{3}z^{-1})(1-\frac{1}{3}z)}$$

then

$$r_y(k) = \tfrac{27}{32}(\tfrac{1}{3})^{|k|}$$

(c) The cross-correlation $r_{xy}(k)$ between $x(n)$ and $y(n)$ is

$$r_{xy}(k) = r_x(k) * h(-k)$$

This may be easily computed using z-transforms as follows,

$$\begin{aligned} P_{xy}(z) &= P_x(z)H(z^{-1}) = \frac{3/4}{(1-\frac{1}{2}z^{-1})(1-\frac{1}{2}z)} \cdot \frac{1-\frac{1}{2}z}{1-\frac{1}{3}z} \\ &= \frac{3/4}{(1-\frac{1}{2}z^{-1})(1-\frac{1}{3}z)} \end{aligned}$$

Writing this in terms of z^{-1} and performing a partial fraction expansion gives

$$P_{xy}(z) = \tfrac{3}{4}\frac{z^{-1}}{(1-\frac{1}{2}z^{-1})(z^{-1}-\frac{1}{3})} = \frac{9/10}{1-\frac{1}{2}z^{-1}} + \frac{3/10}{z^{-1}-\frac{1}{3}}$$

Inverse z-transforming gives

$$r_{xy}(k) = \tfrac{9}{10}(\tfrac{1}{2})^k u(k) + \tfrac{9}{10}(3)^{-k}u(-k-1)$$

(d) The cross-power spectral density, $P_{xy}(z)$, as computed in part (a), is

$$P_{xy}(z) = \frac{3/4}{(1-\frac{1}{2}z^{-1})(1-\frac{1}{3}z)}$$

(e) The cross-correlation, $r_{xy}(k)$, between $x(n)$ and $y(n)$ may found by computing the inverse z-transform of the cross-power spectral density,

$$P_{xy}(z) = \tfrac{3}{4}\frac{z^{-1}}{(1-\frac{1}{2}z^{-1})(z^{-1}-\frac{1}{3})} = \frac{\frac{9}{10}}{1-\frac{1}{2}z^{-1}} + \frac{\frac{3}{10}}{z^{-1}-\frac{1}{3}}$$

Inverse transforming gives

---

4. Hayes Problem 3.6

Solution

(a) Expanding $P_x(e^{j\omega})$ in terms of complex exponentials,

$$P_x(e^{j\omega}) = 3 + 2\cos\omega = 3 + e^{-j\omega} + e^{j\omega}$$

it follows that $r_x(0) = 3$ and $r_x(1) = r_x(-1) = 1$.

(b) Recall the DTFT pair

$$\alpha^{|k|} \longleftrightarrow \frac{1-\alpha^2}{(1-\alpha e^{-j\omega})(1-\alpha e^{j\omega})} = \frac{1-\alpha^2}{(1+\alpha^2) - 2\alpha\cos\omega}$$

Since

$$P_x(e^{j\omega}) = \frac{1}{5+3\cos\omega} = \frac{1/5}{1+\frac{3}{5}\cos\omega}$$

it follows that

$$r_x(k) = \tfrac{1}{4}(-\tfrac{1}{3})^{|k|}$$

(c) With

$$P_x(z) = \frac{-2z^2+5z-2}{3z^2+10z+3} = \frac{-2z+5-2z^{-1}}{(3+z)(3+z^{-1})} = \frac{1}{9}\frac{-2z+5-2z^{-1}}{(1+\frac{1}{3}z)(1+\frac{1}{3}z^{-1})}$$

using the pair

$$(-\tfrac{1}{3})^{|k|} \longleftrightarrow \frac{\frac{8}{9}}{(1+\frac{1}{3}z)(1+\frac{1}{3}z^{-1})}$$

it follows that

$$r_x(k) = \tfrac{5}{8}(-\tfrac{1}{3})^{|k|} - \tfrac{2}{8}(-\tfrac{1}{3})^{|k-1|} - \tfrac{2}{8}(-\tfrac{1}{3})^{|k+1|}$$