EE 514, Fall 2006

Exam 3: Due ECE front desk, 1:45pm, November 16, 2006

Solutions (version: November 9, 2006, 16:58) 75 mins.; Total 50 pts.

1. (10 pts.) Suppose X and Y are jointly Gaussian, where X has mean $m_X = 5$ and variance $\sigma_X^2 = 4$, and Y has mean $m_Y = 3$ and variance $\sigma_Y^2 = 1$. Their correlation coefficient is $\rho = 0.5$. Find the density function of $Z = (X + Y)/2$.

Ans.: Because X and Y are jointly Gaussian, Z is Gaussian. So, to write down its density, we need to compute its mean and variance. It is clear that the mean of Z is $(5 + 3)/2 = 4$. To find the variance of Z, we first write the covariance matrix C of $[X, Y]'$. First note that the covariance of X and Y is $0.5 \cdot \sqrt{4} \cdot \sqrt{1} = 1$. Hence,

$$C = \begin{bmatrix} 4 & 1 \\ 1 & 1 \end{bmatrix}$$

Letting $\boldsymbol{a} = (1/2)[1, 1]$, we can write $Z = \boldsymbol{a}[X, Y]'$, which implies that the variance of Z is given by $\boldsymbol{aCa}' = 7/4$. So the density function of Z is

$$f_Z(z) = \frac{1}{\sqrt{7\pi/2}} \exp\left(-\frac{2(z-4)^2}{7}\right).$$

2. (15 pts.) Suppose we have two real-valued measurements of some physical quantity. Represent these two measurements by zero-mean random variables X and Y. The covariance matrix of $[X, Y]'$ is

$$C = \begin{bmatrix} 6 & 4 \\ 4 & 6 \end{bmatrix}$$

Because of computation and storage limitations, we wish to store only *one* real number, which is a function of X and Y. Our goal is to design a scheme that stores only a single real number and reconstructs the two real-valued measurements as a linear function of the stored number.

a. Design a scheme such that using the single real number to be stored, the mean squared error of the reconstruction is minimized. In other words, if $\hat{X}$ and $\hat{Y}$ are the reconstructed versions of X and Y, then we want to minimize $\mathsf{E}[\|[\hat{X}, \hat{Y}]' - [X, Y]'\|^2]$ (where $\hat{X}$ and $\hat{Y}$ are linear functions of the single real number stored).

b. Write down expressions for $\hat{X}$ and $\hat{Y}$ in terms of X and Y,

c. Write down the resulting mean squared error of the scheme.

Ans.: a. First, we compute eigenvalues and eigenvectors of C so that we can write down the decorrelating transformation. The eigenvalues of C are computed by finding the roots of the characteristic equation $\det(\lambda I - C) = 0$. Plugging in C, we get $\lambda^2 - 12\lambda + 20 = (\lambda - 10)(\lambda - 2) = 0$, which means that the two eigenvalues are $\lambda_1 = 10$ and $\lambda_2 = 2$. The normalized eigenvectors v_1

and v_2 satisfy $(10I - C)v_1 = 0$ and $(2I - C)v_2 = 0$. After some algebra, we get $v_1 = [1, 1]'/\sqrt{2}$ and $v_2 = [1, -1]'/\sqrt{2}$. Hence, the decorrelating transformation is

$$P' = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Let $Z = [Z_1, Z_2]' = P'[X, Y]' = (1/\sqrt{2})[X + Y, X - Y]'$, so that Z_1 and Z_2 are uncorrelated and have variance 10 and 2, respectively. Hence, to minimize the MSE, we store only $Z_1 = (X + Y)/\sqrt{2}$.

b. To reconstruct, we use $[\hat{X}, \hat{Y}]' = Z_1 v_1$, which means that $\hat{X} = \hat{Y} = (X + Y)/2$ (i.e., the average of X and Y).

c. The mean squared error of the scheme is just the variance of Z_2, which is 2.

3. (25 pts.) Suppose we wish to estimate the distance to a target. We model this distance by a Gaussian random variable X with mean 10 and variance 2. To estimate X, we are given two measurements Y_1 and Y_2 such that $Y_i = X + W_i$, $i = 1, 2$, where W_1 and W_2 represent i.i.d. Gaussian noise, independent of X, with mean 0 and variance 1. A simple and intuitive way to estimate X is simply to average the two measurements. In this question, we will explore more rigorous ways of deriving estimators, and see how these estimators are different (if at all) from simple averaging.

For each of the following questions, write the answer as a simple expression involving Y_1 and Y_2. Also, identify the differences (if any) between each estimator and simply averaging the two measurements.

a. Find the Wiener filter for estimating X based on Y_1 and Y_2.

b. Find the ML estimator of X based on Y_1 and Y_2.

c. Find the MAP estimator of X based on Y_1 and Y_2.

d. Find the MMSE estimator of X based on Y_1 and Y_2.

Ans.: a. Let $e = [1, 1]'$, $Y = [Y_1, Y_2]'$, and $W = [W_1, W_2]'$, so that we can write $Y = Xe + W$. We have $m_X = 10$ and $C_X = 2$, and hence $m_Y = 10e$. Then,

$$\begin{aligned} C_{XY} &= \mathsf{E}[(X - 10)(Y - m_Y)'] \\ &= \mathsf{E}[(X - 10)(Xe + W - 10e)'] \\ &= \mathsf{E}[(X - 10)((X - 10)e' + W')] \\ &= \mathsf{E}[(X - 10)^2]e' + \mathsf{E}[(X - 10)W'] \\ &= C_X e' + \mathsf{E}[(X - 10)]\mathsf{E}[W'] \quad \text{by independence} \\ &= 2e'. \end{aligned}$$

Also,

$$\begin{aligned} C_Y &= \mathsf{E}[((X - 10)e + W)((X - 10)e + W)'] \\ &= \mathsf{E}[(X - 10)^2]ee' + \mathsf{E}[WW'] \quad \text{by independence} \\ &= 2ee' + I. \end{aligned}$$

Hence, the Wiener filter is given by $A(Y - m_Y) + m_X = AY + 10(1 - Ae)$, where

$$A = C_{XY}C_Y^{-1} = 2e'(2ee' + I)^{-1} = [2, 2]\begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}^{-1} = \left[\frac{2}{5}, \frac{2}{5}\right] = \frac{2}{5}e'.$$

To summarize, the Wiener filter is given by

$$\frac{2}{5}(Y_1 + Y_2) + 2.$$

(Note: The fact that the noise is Gaussian is irrelevant here.)

Simple averaging corresponds to $(Y_1 + Y_2)/2$. Hence, there are two difference between Wiener filtering and averaging in this case: the first is a multiplicative factor of $4/5$, and the second is an additive term of 2.

b. The conditional CDF of $[Y_1, Y_2]'$ given $X = x$ is

$$\begin{aligned} F_{Y_1,Y_2|X}(y_1, y_2|x) &= \mathsf{P}\{Y_1 \le y_1, Y_2 \le y_2 | X = x\} \\ &= \mathsf{P}\{X + W_1 \le y_1, X + W_2 \le y_2 | X = x\} \\ &= \mathsf{P}\{W_1 \le y_1 - x, W_2 \le y_2 - x | X = x\} \\ &= \mathsf{P}\{W_1 \le y_1 - x, W_2 \le y_2 - x\} \quad \text{by independence} \\ &= \mathsf{P}\{W_1 \le y_1 - x\}\mathsf{P}\{W_2 \le y_2 - x\} \quad \text{by independence} \\ &= F_W(y_1 - x)F_W(y_2 - x), \end{aligned}$$

where F_W is the CDF of W ($N(0, 1)$). Hence, the likelihood function is given by

$$f_{Y_1,Y_2|x}(y_1, y_2|x) = f_W(y_1 - x)f_W(y_2 - x). = \frac{1}{2\pi}\exp(-(1/2)((y_1 - x)^2 + (y_2 - x)^2)).$$

Maximizing with respect to x is equivalent to maximizing the exponent with respect to x, which is equivalent to minimizing the quadratic function $(y_1 - x)^2 + (y_2 - x)^2$. Some simple calculations give $(y_1 + y_2)/2$.

To summarize, the ML estimator is given by

$$\frac{1}{2}(Y_1 + Y_2),$$

which is the same as simple averaging.

c. Because X, W_1, and W_2 are independent, and they are all Gaussian, $[Y_1, Y_2]'$ is a Gaussian random vector. Hence, the MAP estimator is the same as the Wiener filter: $2(Y_1 + Y_2)/5 + 2$. The differences with simple averaging are as in part a.

Aside: The MAP estimator maximizes

$$f_{Y_1,Y_2|x}(y_1, y_2|x)f_X(x) = \text{constant} \cdot \exp(-(1/2)((y_1 - x)^2 + (y_2 - x)^2 + (x - 10)^2/2)),$$

which is the same as minimizing the quadratic $(y_1 - x)^2 + (y_2 - x)^2 + (x - 10)^2/2$. Again, a simple calculation gives the same answer as Wiener filtering.

d. Same answer as part c.