ECE 514, Fall 2007

Exam 3: Due ECE front desk, 3:00pm, December 7, 2007

Solutions (version: December 13, 2007, 8:19) 75 mins.; Total 50 pts.

1. (15 pts.) Suppose we wish to estimate a quantity represented by a real-valued random variable X with mean m and variance σ_X^2. For this estimation, we take n measurements of X represented by $Y_i = X + W_i$, $i = 1, \ldots, n$, where $W_1, \ldots, W_n$ represent i.i.d. noise, independent of X, with mean 0 and variance σ_W^2.

a. Let $e = [1, \ldots, 1]'$. Show that

$$e'(\sigma_X^2 ee' + \sigma_W^2 I)^{-1} = \frac{1}{n\sigma_X^2 + \sigma_W^2} e'.$$

b. Find the Wiener filter for estimating X based on $Y_1, \ldots, Y_n$. (Part a helps.)

c. What exactly is the difference between the filter in part b and simply averaging the measurements? What is the effect of very large n, or very large σ_X^2, or very large σ_W^2?

Ans.: a. Easy to show by verifying that

$$e'(\sigma_X^2 ee' + \sigma_W^2 I) = (n\sigma_X^2 + \sigma_W^2)e',$$

using the identity $e'e = n$. Note that this identity states that the matrix $(\sigma_X^2 ee' + \sigma_W^2 I)$ has e as an eigenvector and $(n\sigma_X^2 + \sigma_W^2)$ as its corresponding eigenvalue.

b. Let $W = [W_1, \ldots, W_n]'$, so that we can write $Y = Xe + W$. We have $m_X = m$ and $C_X = \sigma_X^2$, and hence $m_Y = me$. Then,

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

Also,

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

Hence, writing $Y = [Y_1, \ldots, Y_n]'$, the Wiener filter is given by $A(Y - m_Y) + m_X = AY + m(1 - Ae)$, where

$$A = C_{XY}C_Y^{-1} = \sigma_X^2 e'(\sigma_X^2 ee' + \sigma_W^2 I)^{-1} = \frac{\sigma_X^2}{n\sigma_X^2 + \sigma_W^2} e'$$

by part a. Hence, the filter is given by

$$\frac{\sigma_X^2}{n\sigma_X^2 + \sigma_W^2} e'Y + m\left(1 - \frac{n\sigma_X^2}{n\sigma_X^2 + \sigma_W^2}\right).$$

c. We can rewrite the filter as

$$\frac{n\sigma_X^2}{n\sigma_X^2 + \sigma_W^2} \frac{e'Y}{n} + \frac{m\sigma_W^2}{n\sigma_X^2 + \sigma_W^2}.$$

Averaging the measurements corresponds to

$$\frac{e'Y}{n}.$$

Hence, there are two differences: There is a multiplicative factor of

$$\frac{n\sigma_X^2}{n\sigma_X^2 + \sigma_W^2}$$

and an additive term of

$$\frac{m\sigma_W^2}{n\sigma_X^2 + \sigma_W^2}.$$

If n is very large, the multiplicative factor is close to 1, and the additive term is close to 0. In this case, the Wiener filter is close to simple averaging.

If σ_X^2 is very large, then again the multiplicative factor is close to 1, and the additive term is close to 0. In this case, the Wiener filter is close to simple averaging.

If σ_W^2 is very large, then the multiplicative factor is close to $n\sigma_X^2/\sigma_W^2$ (a small number), and the additive term is close to m. In this case, the Wiener filter discounts the averaging of the measurements, and simply uses m as an estimate of X. This makes sense.

2. (10 pts.) Suppose we pass an nth-order strictly stationary random process through a linear time invariant system. Is the output necessarily nth-order stationary? Explain fully. (For your analysis you may assume either discrete or continuous time.)

Ans.: The answer is no. To see why, consider a first-order strictly stationary process (in discrete time) $X_1, X_2, \ldots$ for which the distribution of $X_1 + X_2$ is not equal to the distribution of $X_3 + X_4$. Such a process exists (see below for a specific example)—the distribution of the sum $X_t + X_{t+1}$ depends on the joint distribution of (X_t, X_{t+1}), which in general depends on t for a first-order stationary process that is not second-order stationary. So, if we pass the process X_t through an LTI system whose output is given by $Y_t = X_t + X_{t+1}$, then the distribution of Y_1 is not equal to the distribution of Y_3 (i.e., Y_t is not first-order strictly stationary).

A specific example of a process meeting the above requirement is as follows. Consider a sequence of random variables $X_1, X_2, \ldots$ such that each X_t is $N(0, 1)$ (hence the process is first-order stationary). Suppose that $X_1 = X_2$ but X_3 and X_4 are independent. Then, the variance of $X_1 + X_2$ is 4, but the variance of $X_3 + X_4$ is only 2.

3. (10 pts.) Consider the setting of the matched-filter problem (Section 10.7), where $v(t)$ is a given real signal such that $V(f) = 1$ for $|f| \in [1, 2]$ and $V(f) = 0$ elsewhere. Suppose you treat the random process X_t as a jamming signal that you are free to design (as an adversary). The design objective is to minimize the best-case SNR at the output of the receiver (*best-case* here means maximum over all possible receiver designs). The design constraint is that you have only a fixed amount of power P_0 to allocate to your jammer. Describe your design of the jamming signal. (It suffices to design the power spectral density S_X.)

Ans.: We know that the best-case output SNR is given by the formula

$$\int_{-\infty}^{\infty} \frac{|V(f)|^2}{S_X(f)}\, df = 2\int_1^2 \frac{1}{S_X(f)}\, df.$$

We wish to minimize this expression by choice of S_X, subject to an upper-bound constraint on the jamming power: $\int_{-\infty}^{\infty} S_X(f)\, df \le P_0$. It is clear that we should allocate all of the power P_0, so we can treat this inequality constraint as an equality constraint.

For the design of S_X, we first impose the condition that $S_X(f) = 0$ if $|f| \notin [1, 2]$ (i.e., we shouldn't allocate any power in those frequencies where the signal v is not present). So the problem is to find an S_X to minimize

$$\int_1^2 \frac{1}{S_X(f)}\, df$$

subject to

$$\int_1^2 S_X(f)\, df = \frac{P_0}{2}.$$

What positive function has a given integral but has minimal integral of its reciprocal? The intuitive answer is the constant function. To see this, consider a nonconstant function; say $S_X(f_1) = a$ and $S_X(f_2) = b \neq a$. Consider the infinitesimal area $1/S_X(f_1)\, df + 1/S_X(f_2)\, df = (1/a + 1/b)\, df$. Now, we can reallocate power between f_1 and f_2 in a way that doesn't change the sum at these two frequencies, to get a new PSD S'_X: $S'_X(f_1) = S'_X(f_2) = (a + b)/2$. However, notice that $1/S'_X(f_1)\, df + 1/S'_X(f_2)\, df = (2/(a + b))\, df < (1/a + 1/b)\, df$.

So the design is this: $S_X(f) = P_0/2$ for $|f| \in [1, 2]$ and $S_X(f) = 0$ elsewhere.

4. (15 pts.) Let B be a Bernoulli random variable taking values on $\{0, 1\}$. Define the discrete-time process $X_1, X_2, \ldots$ by $X_t = (-1)^{B+t}$, $t = 1, 2, \ldots$. Let $p = \mathsf{P}\{B = 0\}$.

a. For what values of p is the process X_t strictly stationary?

b. For what values of p in part a is the process also ergodic?

Explain fully.

Ans.: a. If $p \neq 1/2$, the process is definitely not stationary. In particular,

$$\mathsf{E}[X_t] = \mathsf{E}[(-1)^{B+t}] = p(-1)^t + (1 - p)(-1)^{t+1} = (2p - 1)(-1)^t,$$

which depends on t if $p \neq 1/2$.

If $p = 1/2$, the process is indeed stationary. To see this, note that the law of $X_t = (-1)^{B+t}$ depends only on the distribution of B. Now, consider a time shift τ. If τ is even, then

$$X_{t+\tau} = (-1)^{B+t+\tau} = (-1)^{\tau} \cdot (-1)^{B+t} = (-1)^{B+t} = X_t.$$

In this case, the law of $X_{t+\tau}$ is clearly equal to that of X_t. If τ is odd, then

$$X_{t+\tau} = (-1)^{B+t+\tau} = (-1)^{\tau}(-1)^{B+t} = (-1) \cdot (-1)^{B+t} = -X_t = (-1)^{B'+t}$$

where $B' = 1 - B$ (i.e., for B' the roles of 1 and 0 are reversed relative to B). But if $p = 1/2$, the distribution of B' is the same as that of B. Hence, the law of $X_{t+\tau}$ is equal to that of X_t. Hence, we have shown that for any τ, the law of $X_{t+\tau}$ is equal to that of X_t.

b. The only value of p in part a is $1/2$. The question is whether or not for $p = 1/2$ the process is ergodic. The answer is yes. To see this, let A be a T-invariant set. For convenience, let $\boldsymbol{x}$ be the sequence $(-1)^t$ and $\boldsymbol{X}$ the process $X_1, X_2, \ldots$. Notice that the process $\boldsymbol{X}$ has only two possible sample paths (a.s.): $\boldsymbol{x}$ and $-\boldsymbol{x}$. Notice also that $T\boldsymbol{x} = -\boldsymbol{x}$.

It remains to show that $\mathsf{P}\{\boldsymbol{X} \in A\} = 0$ or 1. If $\boldsymbol{x} \notin A$, then clearly $\mathsf{P}\{\boldsymbol{X} \in A\} = 0$. If $\boldsymbol{x} \in A$ then $-\boldsymbol{x} = T\boldsymbol{x} \in A$ also, because A is T-invariant; so, in this case, $\mathsf{P}\{\boldsymbol{X} \in A\} = 1$. Hence, $\boldsymbol{X}$ is ergodic.