16:332:541 **Stochastic Signals and Systems** **Midterm Exam**

November 8, 2006

Solution

You have 90 minutes to complete the first three problems of this exam. You are invited to complete Problem 4 at home and to submit your solution in class on Monday. *The take home component must be completed alone without collaboration or assistance from other people.* Items with unspecified point values are worth ten points. Please read both sides of the exam carefully and ask the instructor if you have any questions.

Preliminary: *(10 points)* Put your name and your Rutgers netid on the front of each exam bluebook. Make up an ostensibly random (but personally memorable) four digit code. Write this code on the *upper right corner of the inside front cover* of your first bluebook. This code must not contain any part of your SSN or Rutgers ID number. (This code will be used to post grades at the end of the semester.

1. *40 points* Random variables X_1 and X_2 have zero expected value. The random vector $\mathbf{X} = \begin{bmatrix} X_1 & X_2 \end{bmatrix}'$ has a covariance matrix of the form

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

   (a) For what values of α and β is $\mathbf{C}$ a valid covariance matrix?

   $\mathbf{C}$ *must be symmetric since*

$$\alpha = \beta = E[X_1 X_2].$$

   In addition, α must be chosen so that $\mathbf{C}$ *is positive semi-definite. Since the characteristic equation is*

$$det(\mathbf{C} - \lambda\mathbf{I}) = (1-\lambda)(4-\lambda) - \alpha^2 = \lambda^2 - 5\lambda + 4 - \alpha^2 = 0,$$

   the eigenvalues of $\mathbf{C}$ *are*

$$\lambda_{1,2} = \frac{5 \pm \sqrt{25 - 4(4-\alpha^2)}}{2}.$$

   The eigenvalues are non-negative as long as $\alpha^2 \leq 4$, or $|\alpha| \leq 2$. Another way to reach this conclusion is through the requirement that $|\rho_{X_1 X_2}| \leq 1$.

   (b) For what values of α and β can $\mathbf{X}$ be a Gaussian random vector?

   It remains true that $\alpha = \beta$ and $\mathbf{C}$ *must be positive semi-definite. For* $\mathbf{X}$ *to be a Gaussian vector,* $\mathbf{C}$ *also must be positive definite. For the eigenvalues of* $\mathbf{C}$ *to be strictly positive, we must have $|\alpha| < 2$.*

   (c) *20 points* Suppose now that α and β satisfy the conditions in part (b) and $\mathbf{X}$ is a Gaussian random vector.

      i. What is the PDF of X_2?

      From the problem statement and the matrix $\mathbf{C}$*, X_2 is a Gaussian $(0, \sigma = 2)$ random variable and thus has PDF*

$$f_{X_2}(x) = \frac{1}{\sqrt{8\pi}} e^{-x^2/8}.$$

      ii. What is the PDF of $W = 2X_1 - X_2$?

Since $\mathbf{X}$ *is a Gaussian vector,* W *is a Gaussian random variable. Thus, we need only calculate*

$$E[W] = 2E[X_1] - E[X_2] = 0,$$

and

$$\begin{aligned}\sigma_W^2 = E[W^2] &= E[4X_1^2 - 4X_1X_2 + X_2^2] \\ &= 4\sigma_{X_1}^2 - 4cov[X_1, X_2] + \sigma_{X_2}^2 \\ &= 4 - 4\alpha + 4 = 4(2-\alpha).\end{aligned}$$

The PDF of W *is*

$$f_W(w) = \frac{1}{\sqrt{8(2-\alpha)\pi}} e^{-w^2/8(2-\alpha)}.$$

2. *60 points* Starting infinitely long ago in the past, a new customer arrives each minute at a bank, exactly at the start of each minute. Each arriving customer is immediately served by a teller. (There are always as many tellers as needed to serve all customers in the bank.) After each minute of service, a customer departs with probability $1-p$, independent of the departures of all other customers. Departures at the end of minute $t-1$ occur the instant before the new customer arrives at the start of minute t. We say that a customer is in service at minute t, if the customer arrived at a minute $s \leq t$ and did not depart prior to the end of minute t.

   (a) Let $X_{t,k}(t)$ denote an indicator random variable such that $X_{t,k} = 1$ if the customer that arrived at the start of minute t is still in service at minute $t+k$; otherwise $X_{t,k} = 0$. What is the probability mass function (PMF) $P_{X_{t,k}}(x)$ of $X_{t,k}$?
   The customer that arrived at the start of minute t *stays in service at the end of each minute with probability* p. *This customer is in service at time* $t+k$ *with probability* p^k. *In this case,* $X_{t,k} = 1$; *otherwise* $X_{t,k} = 0$. *The Bernoulli PMF is*

$$P_{X_{t,k}}(x) = \begin{cases} p^k & x = 1, \\ 1 - p^k & x = 0, \\ 0 & \textit{otherwise.} \end{cases}$$

   (b) *20 points* Let Y_t denote the number of customers in service at minute t (i.e. the instant after the new arrival at the start of minute t.) Find the expected value $E[Y_t]$ and variance $\sigma_{Y_t}^2$. Hint: $Y_t = X_{t,0} + X_{t-1,1} + \cdots$.
   We will use the fairly huge hint. Since $E[X_{t,k}] = P\{X_{t,k} = 1\} = p^k$,

$$E[Y_t] = E\left[\sum_{k=0}^{\infty} X_{t-k,k}\right] = \sum_{k=0}^{\infty} E[X_{t-k,k}] = \sum_{k=0}^{\infty} p^k = \frac{1}{1-p}.$$

   Since a Bernoulli (p) *random variable has variance* $p(1-p)$, $\sigma_{X_{t,k}}^2 = p^k(1-p^k)$. *Next we observe for* $k \neq k'$ *that* $X_{t-k,k}$ *and* $X_{t-k',k'}$ *are independent since the service time of a customer arriving at time* $t-k$ *is independent of the service time of the customer arriving at time* $t-k'$. *Since the variance of a sum of independent random variables equals the sum of the variances,*

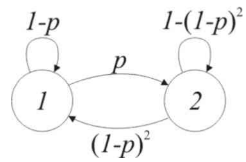
$$\begin{aligned}\sigma_{Y_t}^2 = \sum_{k=0}^{\infty} \sigma_{X_{t-k,k}}^2 &= \sum_{k=0}^{\infty} p^k(1-p^k) \\ &= \sum_{k=0}^{\infty} p^k - \sum_{k=0}^{\infty} (p^2)^k = \frac{1}{1-p} - \frac{1}{1-p^2} = \frac{p}{1-p^2}.\end{aligned}$$

(c) *30 points* Suppose that the customers are afraid of crowded places. Upon arrival, a customer instantly departs if the bank already has two customers. Model the process Y_t as a discrete-time Markov chain. Either find the limiting state probabilities or explain why they do not exist.

Since Y_t is the number of customers just after the new arrival, we know that $Y_t \geq 1$. Since the customers don't like to be crowded, $Y_t \leq 2$. Thus the Markov chain describing Y_t has state space $\{1, 2\}$. The state transitions can be found from the following observations:

- *If $Y_t = 1$, then $Y_{t+1} = 1$ if and only if the customer in service departs, which occurs with probability $p_{11} = 1 - p$.*
- *If $Y_t = 2$, then $Y_{t+1} = 1$ if and only if both customers in service depart. This occurs with probability $p_{21} = (1-p)^2$.*

The two-state Markov chain is



The chain is (obviously) irreducible and positive recurrent. The stationary probabilities satisfy $\pi_1 p = \pi_2 (1-p)^2$. This implies

$$\pi_1 = \frac{(1-p)^2}{(1-p)^2 + p}, \qquad \pi_2 = \frac{p}{(1-p)^2 + p}.$$

3. *70 points* At time $t = 0$, the price of a stock is a constant k dollars. At time $t > 0$, the price X of the stock is a uniform $(k-t, k+t)$ random variable. At time t, a *Put Option at Strike* k (which is the right to sell the stock at price k) has value $V = (k - X)^+$ where the operator $(\cdot)^+$ is defined as $(z)^+ = \max(z, 0)$. You may also recall that a *Call Option at Strike* k (the right to buy the stock at price k) has value $W = (X - k)^+$.

(a) *20 points* At time 0, you sell the put and receive d dollars. At time t, you purchase the put for V dollars to cancel your position. Your profit is $R = d - V$. Find the central moments $E[R]$ and σ_R^2.

Since $R = d - V$,

$$E[R] = d - E[V], \qquad \sigma_R^2 = \sigma_V^2.$$

From the problem statement, the PDF of X is

$$f_X(x) = \begin{cases} 1/(2t) & k - t \leq x \leq k + t, \\ 0 & \textit{otherwise.} \end{cases}$$

It follows that

$$\begin{aligned} E[V] = E\left[(k-X)^+\right] &= \int_{-\infty}^{\infty} (k-x)^+ f_X(x)\, dx \\ &= \int_{k-t}^{k+t} (k-x)^+ \frac{1}{2t}\, dx \\ &= \frac{1}{2t} \int_{k-t}^{k} (k-x)\, dx = -\frac{1}{4t}(k-x)^2 \bigg|_{x=k-t}^{x=k} = \frac{t}{4}. \end{aligned}$$

and

$$E\big[V^2\big] = E\big[((k-X)^+)^2\big] = \int_{-\infty}^{\infty} ((k-x)^+)^2 f_X(x)\,dx$$
$$= \int_{k-t}^{k+t} ((k-x)^+)^2 \frac{1}{2t}\,dx$$
$$= \frac{1}{2t}\int_{k-t}^{k} (k-x)^2\,dx = -\frac{1}{6t}(k-x)^3\Big|_{x=k-t}^{x=k} = \frac{t^2}{6}.$$

It follows that

$$\sigma_V^2 = E\big[V^2\big] - (E[V])^2 = \frac{t^2}{6} - \frac{t^2}{16} = \frac{5t^2}{48}.$$

Finally, we go back to the beginning to write

$$E[R] = d - \frac{t}{4}, \qquad \sigma_R^2 = \sigma_V^2 = \frac{5t^2}{48}.$$

(b) *20 points* In a short straddle, you sell the put for d dollars and you also sell the call for d dollars. At time t, you purchase the put for V dollars and the call for W dollars to cancel both positions. Your profit is

$$R' = 2d - (V + W).$$

Find the expected value $E[R']$ and variance $\sigma_{R'}^2$.
Since the PDF of X is symmetric around $x = k$, one can deduce from symmetry that V and W are identically distributed. Thus $E[W] = E[V] = t/4$ and $E\big[W^2\big] = E\big[V^2\big] = t^2/6$. Thus,

$$E[V+W] = 2E[V] = \frac{t}{2}$$

and

$$E\big[R'\big] = 2d - E[V+W] = 2d - \frac{t}{2} = 2E[R].$$

Since $R' = 2d - (V+W)$, $\sigma_{R'}^2 = \sigma_{V+W}^2$. This requires us to find $E\big[(V+W)^2\big]$. At is point, we note that V and W are not independent. In fact, if $V > 0$, then $W = 0$ but if $W > 0$, then $V = 0$. This implies that $VW = 0$ and, of course, $E[VW] = 0$. It then follows that

$$E\big[(V+W)^2\big] = E\big[V^2 + 2VW + W^2\big] = E\big[V^2\big] + E\big[W^2\big] = 2E\big[V^2\big] = \frac{t^2}{3}.$$

Finally, we can write

$$\sigma_{R'}^2 = \sigma_{V+W}^2 = E\big[(V+W)^2\big] - (E[V+W])^2 = \frac{t^2}{3} - \left(\frac{t}{2}\right)^2 = \frac{t^2}{12}.$$

(c) *20 points* Find the PDF $f_{R'}(r)$ of R'.
First we find the CDF

$$F_{R'}(r) = P\big\{R' \le r\big\}P\{2d - (V+W) \le r\} = P\{V + W \ge 2d - r\}.$$

Since $V + W$ is non-negative,

$$F_{R'}(r) = P\{V + W \ge 2d - r\} = 1, \qquad r \ge 2d.$$

Now we focus on the case $r \le 2d$. Here we observe that $V > 0$ and $W > 0$ are mutually exclusive events. Thus, for $2d - r \ge 0$,

$$F_{R'}(r) = P\{V \ge 2d - r\} + P\{W \ge 2d - r\} = 2P\{W \ge 2d - r\}$$

since W and V are identically distributed. Since $W = (X - k)^+$ and $2d - r \ge 0$,

$$P\{W \ge 2d - r\} = P\big\{(X - k)^+ \ge 2d - r\big\} = P\{X - k \ge 2d - r\}$$
$$= \begin{cases} 0 & (2d - r) > t, \\ \frac{t-(2d-r)}{2t} & (2d - r) \le t. \end{cases}$$

We can combine the above results in the following statement:

$$F_{R'}(r) = 2P\{W \ge 2d - r\} = \begin{cases} 0 & r < 2d - t, \\ \frac{t-2d+r}{t} & 2d - t \le r \le 2d, \\ 1 & r \ge 2d. \end{cases}$$

The PDF of R' is

$$f_{R'}(r) = \begin{cases} \frac{1}{t} & 2d - t \le r \le 2d, \\ 0 & \text{otherwise}. \end{cases}$$

(d) *10 points* Suppose d is sufficiently large that $E[R'] > 0$. Would you be interested in selling the short straddle? Are you getting something, namely $E[R']$ dollars, for nothing?
A general answer is that you may expect to receive a return of $E[R'] > 0$ dollars; however this is not free because you assume the risk of a significant loss. In a real investment, the PDF of the price X is not bounded and the loss can be very very large. However, in the case of this problem, the bounded PDF for X implies the loss is not so terrible. From part (a), or by examination of the PDF $f_{R'}(r)$, we see that

$$E[R'] = \frac{4d - t}{2}.$$

Thus $E[R'] > 0$ if and only if $d > t/4$. In the worst case of $d = t/4$, we observe that R' has a uniform PDF over $(-t/2, t/2)$ and the worst possible loss is $t/2$ dollars. Whether the risk of such a loss is worth taking for an expected return $E[R']$ would depend mostly on your financial capital and your investment objectives, which were not indcluded in the problem formulation.

4. *70 points* Random variable $Y = X - Z$ is a noisy observation of the continuous random variable X. In particular, the noise Z has zero mean and unit variance and is independent of X. We wish to use Y to estimate X by forming an estimator of the form $\hat{X} = \hat{X}(Y) = aY$.

(a) As a function of a and perhaps the moments of X, find the mean square error $e(a) = E\Big[(X - \hat{X})^2\Big]$.
Since $\hat{X} = a(X - Z)$,

$$\begin{aligned} e(a) &= E\Big[(X - a(X - Z))^2\Big] \\ &= E\Big[((1 - a)X + aZ)^2\Big] \\ &= E\big[(1 - a)^2X^2 + a(1 - a)XZ + a^2Z^2\big]. \end{aligned}$$

Since, $E[Z] = 0$, $E\big[Z^2\big] = \sigma_Z^2 = 1$. Also, since X and Z are independent, $E[XZ] = E[X]E[Z] = 0$. These facts imply

$$e(a) = (1 - a)^2E\big[X^2\big] + a^2E\big[Z^2\big] = (1 - a)^2E\big[X^2\big] + a^2.$$

(b) Find $\hat{a}$, the value of a that minimizes the mean square estimation error $e(a)$.

We choose $\hat{a}$ *such that*

$$\left.\frac{de(a)}{da}\right|_{a=\hat{a}} = -2(1-\hat{a})E\left[X^2\right] + 2\hat{a} = 0.$$

This implies

$$\hat{a} = \frac{E\left[X^2\right]}{1+E[X^2]}.$$

(c) Let b be an arbitrary constant. Show that the PDF of $W = b+Z$ is $f_W(w) = f_Z(w-b)$.

This is simple. We start with the CDF

$$F_W(w) = P\{W \le w\} = P\{b+Z \le w\} = P\{Z \le \; w-b\} = F_Z(w-b)\,.$$

Taking derivatives with respect to w*, we obtain*

$$f_W(w) = f_Z(w-b)\,.$$

(d) *20 points* Find the conditional expectation $E[X|Y]$.

This part probably should have followed the next part (although this is essentially irrelevant since this question was done at home. Here we need to find

$$E[X|Y=y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\, dx.$$

Replacing y *by* Y *in* $E[X|Y=y]$ *will yield the desired answer. Given* $X=x$, $Y = x-Z$ *and following steps similar to the previous part to write*

$$\begin{aligned} P\{Y \le y|X=x\} &= P\{x - Z \le y|X=x\} \\ &= P\{Z \ge x-y|X=x\} = 1 - F_Z(x-y)\,. \end{aligned}$$

Note the last inequality follows because Z *and* X *are independent random variables. Taking derivatives, we have*

$$f_{Y|X}(y|x) = \frac{dP\{Z \le x-y|X=x\}}{dy} = \frac{d}{dy}\left(1 - F_Z(x-y)\right) = f_Z(x-y)\,.$$

It follows that X *and* Y *have joint PDF*

$$f_{X,Y}(x,y) = f_{Y|X}(y|x)\, f_X(x) = f_Z(x-y)\, f_X(x)\,.$$

By the definition of conditional PDF,

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{f_Z(x-y)\, f_X(x)}{f_Y(y)},$$

and thus

$$E[X|Y=y] = \int_{-\infty}^{\infty} x\frac{f_Z(x-y)\, f_X(x)}{f_Y(y)}\, dx = \frac{1}{f_Y(y)}\int_{-\infty}^{\infty} x f_Z(x-y)\, f_X(x)\, dx.$$

Without more information, this is the simplest possible answer. Also note that the denominator $f_Y(y)$ *is given by*

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx = \int_{-\infty}^{\infty} f_Z(x-y)\, f_X(x)\, dx.$$

For a given PDF $f_Z(z)$*, it is sometimes possible to compute these integrals in closed form; Gaussian* Z *is one such example.*

(e) *20 points* It is known that $X^*(Y) = E[X|Y]$ is the minimum mean square error estimator of X given Y. Consider the following argument:

> *Since $X = Y + Z$, we see that if $Y = y$, then $X = y + Z$. Thus, from the result of part (c) with $W = X$ and $b = y$, the conditional PDF of X given $Y = y$ is $f_{X|Y}(x|y) = f_Z(x - y)$. It follows that*
>
> $$E[X|Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\, dx = \int_{-\infty}^{\infty} x f_Z(x - y)\, dx.$$
>
> *With the variable substitution, $z = x - y$, we obtain*
>
> $$E[X|Y = y] = \int_{-\infty}^{\infty} (z + y) f_Z(z)\, dz = E[Z] + y = y.$$
>
> *We conclude that $E[X|Y] = Y$. Since $E[X|Y]$ is optimal in the mean square square sense, we conclude that the optimal estimator of the form $\hat{X}(Y) = \hat{a}Y$ must satisfy $\hat{a} = 1$.*

Show that the above conclusion is inconsistent with your prior results. What is the error in the above argument?

Note that this question does not *ask you to show that $E[X|Y]$ is the MMSE estimator. This is shown in the text and in many other texts. However, if the best nonlinear estimator $E[X|Y]$ is a linear function $\tilde{X} = \tilde{a}Y$, then the best linear estimator $\hat{a}Y$ must have $\hat{a} = \tilde{a}$. However, we have a contradiction since part (a) showed that the best estimator of the form $\hat{X} = aY$ has an optimal value of $a = \hat{a} < 1$. In particular, the estimator $\tilde{Y} = \tilde{a}Y = Y$ has mean square error $e(1) = 1$ while the optimal linear estimator $\hat{X} = \hat{a}Y$ can be shown to have mean square error $e(\hat{a}) = \hat{a} < 1$. Thus there is definitely an error in the logic of the given argument.*

The goal of this question was to identify this logical error. It is true that given $Y = y$, then $X = y + Z$ and that

$$\begin{aligned} P\{X \le x|Y = y\} &= P\{Y + Z \le x|Y = y\} \\ &= P\{y + Z \le x|Y = y\} \\ &= P\{Z \le x - y|Y = y\}. \end{aligned}$$

Many students then used the followed (mistaken) logic that Z and Y are independent to conclude that $P\{Z \le x - y|Y = y\} = P\{Z \le x - y\} = F_Z(x - y)$. Taking a derivative with respect to x leads to the false conclusion that $f_{X|Y}(x|y) = f_Z(x - y)$.

In fact, Z and Y are dependent. *If this is not obvious consider the special case when $E[X] = 0$. In this case, Y and Z are correlated (and thus dependent) since*

$$cov[Y, Z] = E[YZ] = E[(X - Z)Z] = E[XZ] - E\left[Z^2\right] = -E\left[Z^2\right] < 0.$$

As a result, we cannot use part (c) in that $f_{X|Y}(x|y) \ne f_Z(x - y)$.