## Stochastic Processes

## Midterm 1

## Solutions

1. (15 points)
$\left(\Rightarrow\right.$ ( 7 points)) Let $X_{1}, \cdots, X_{n}$ be jointly Gaussian RVs. By definition, we know

$$
Y=\sum_{i=1}^{n} a_{i} X_{i}
$$

is a Gaussian RV for any real $a_{i}$. So, the case that $a_{1}=1$ and all other $a_{i}=0$, for $i \neq 2$, falls in that category, and we conclude that $Y=X_{1}$ is a Gaussian random variable. Similar argument applies to other $X_{i}$.
$(\Leftarrow(8$ points $))$ Conversely, suppose $X_{1}, X_{2}, \cdots, X_{n}$ are individually Gaussian. This does not necessarily imply that $X_{1}, X_{2}, \cdots, X_{n}$ are jointly Gaussian. This can be explained by the example I mentioned in class. You can find that example in problem 4 of last year's midterm that I've posted on the course web.
2. $(10+10=20$ points $)$

Let $X$ and $Y$ be i.i.d. standard Gaussian random variables.
(a) This problem essentially requires you to prove the following two parts. The first part needs you to explain that $X+Y$ and $X-Y$ are jointly Gaussian, which is not difficult. (This part takes weight of 4 points.)
The second part needs you to use the fact that "jointly Gaussian random variables are uncorrelated if and only if they are independent." Thus, we can check the uncorrelatedness between between $X+Y$ and $X-Y$, i.e. check whether $E[(X+Y)(X-Y)]=E[X+Y] E[X-Y]$.
For that, we compute

$$
E[(X+Y)(X-Y)]=E\left[X^{2}\right]-E\left[Y^{2}\right]=0 . \quad \text { (since } X \text { and } Y \text { are i.i.d.) }
$$

And

$$
E[X+Y] E[X-Y]=(E[X]+E[Y])(E[X]-E[Y])=0 .
$$

Thus, we conclude that $X+Y$ and $X-Y$ are independent.
(b) Find $E\left[X^{3}-Y^{3} \mid X-Y\right]$. (Hint: Use part (a).)

The key to solve this problem is to express $X^{3}-Y^{3}$ in terms of $X+Y$ and $X-Y$ only. It follows

$$
\begin{aligned}
X^{3}-Y^{3} & =(X-Y)\left(X^{2}+X Y+Y^{2}\right) \\
& =(X-Y)\left((X+Y)^{2}-X Y\right) \\
& =(X-Y)\left((X+Y)^{2}-\left((X+Y)^{2}-(X-Y)^{2}\right) / 4\right) \\
& =\frac{3}{4}(X-Y)(X+Y)^{2}+\frac{1}{4}(X-Y)^{3} .
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
E\left[X^{3}-Y^{3} \mid X-Y\right] & =E\left[\left.\frac{3}{4}(X-Y)(X+Y)^{2}+\frac{1}{4}(X-Y)^{3} \right\rvert\, X-Y\right] \\
& =\frac{3}{4}(X-Y) E\left[(X+Y)^{2} \mid X-Y\right]+\frac{1}{4}(X-Y)^{3} \\
& =\frac{3}{4}(X-Y) E\left[(X+Y)^{2}\right]+\frac{1}{4}(X-Y)^{3}
\end{aligned}
$$

(the condition is removed since $X+Y$ and $X-Y$ are indep.)
The 2nd equality, where $X-Y$ is regarded as a constant and moved outside the expectation due to the conditioning on it, can be justified by Prob. 2 of homework 1 . So it remains to find $E\left[(X+Y)^{2}\right]$, which is

$$
\begin{aligned}
E\left[(X+Y)^{2}\right] & =E\left[X^{2}+2 X Y+Y^{2}\right] \\
& =E\left[X^{2}\right]+E\left[Y^{2}\right] \\
& =2 . \quad(X, Y \text { are i.i.d. standard Gaussian })
\end{aligned}
$$

It then gives

$$
E\left[X^{3}-Y^{3} \mid X-Y\right]=\frac{3}{2}(X-Y)+\frac{1}{4}(X-Y)^{3} .
$$

3. ( 10 points $\times 5=50$ points)
(a) We see that $X+\alpha Z$ and $Z$ are jointly Gaussian. Thus, we can find $\alpha$ that makes them uncorrelated, $E[(X+\alpha Z) Z]=0$, to achieve the goal. So we have

$$
E[X Z]+\alpha E\left[Z^{2}\right]=0
$$

From the mean vector and covariance matrix, we know $E\left[Z^{2}\right]=1$ and $E[(X-1) Z]=1$, which leads to $E[X Z]=1$.
Therefore, $\alpha=-1$.
(b) Since $X, Y, Z$ are jointly Gaussian, by definition there any linear combination is a Gaussian RV. So we know $S$ is Gaussian. It's mean is $E[S]=E[X+Y+Z]=3$, and variance is

$$
\begin{aligned}
\operatorname{Var}(S)= & E\left[(S-E[S])^{2}\right] \\
= & E\left[((X-1)+(Y-2)+Z)^{2}\right] \\
= & E\left[(X-1)^{2}\right]+E\left[(Y-2)^{2}\right]+E\left[Z^{2}\right]+2 E[(X-1)(Y-2)] \\
& +2 E[(X-1) Z]+2 E[(Y-2) Z] \\
= & 3+3+1+2+2 \\
= & 11
\end{aligned}
$$

With mean and variance known, it's an easy task to write down the pdf.
(c) Let $\mathbf{r}=[Y, Z]^{T}$. The conditional variance can be found by the formula

$$
\operatorname{Var}(X \mid Y, Z)=\mathbf{K}_{x \mid r}=\mathbf{K}_{x}-\mathbf{K}_{x r} \mathbf{K}_{r}^{-1} \mathbf{K}_{r x}
$$

where $\mathbf{K}_{x}$ is just the variance of $X, \mathbf{K}_{x r}$ is the cross covariance between $X$ and $\mathbf{r}$, and $\mathbf{K}_{r}$ is the covariance matrix of $\mathbf{r}$. Further calculations give

$$
\begin{aligned}
\mathbf{K}_{x r} & =E[(X-1)[Y-2, Z]]=[E[(X-1)(Y-2)], E[(X-1) Z]]=[1,1] \\
\mathbf{K}_{r} & =E\left[[Y-2, Z]^{T}[Y-2, Z]\right] \\
& =\left[\begin{array}{ll}
3 & 0 \\
0 & 1
\end{array}\right]
\end{aligned}
$$

It follows that the conditional variance

$$
\operatorname{Var}(X \mid Y, Z)=3-[1,1]\left[\begin{array}{cc}
\frac{1}{3} & 0 \\
0 & 1
\end{array}\right][1,1]^{T}=5 / 3
$$

(d) This one is straightforward by applying the formula given on the cover page.

$$
\rho_{x y}=E[(X-E[X])(Y-E[Y])] / \sigma_{X} \sigma_{Y}=1 / 3
$$

(e) From the covariance matrix of $\mathbf{w}$ we know $Y, Z$ are independent. Thus we can simply choose the matrix $\mathbf{A}$ to be a diagonal matrix

$$
\mathbf{A}=\left[\begin{array}{cc}
\frac{1}{\sqrt{3}} & 0 \\
0 & 1
\end{array}\right]
$$

such that both the covariance matrix of $\mathbf{A r}$ and $\mathbf{A r}+\mathbf{b}$ for any $\mathbf{b}$, i.e. $\mathbf{A K}_{r} \mathbf{A}^{T}$ is an identity matrix.
Note that adding a vector $\mathbf{b}$ to $\mathbf{A r}$ does NOT alter the covariance matrix. But the mean after adding the vector $\mathbf{b}$ to $\mathbf{A r}$ becomes $\mathbf{A m}_{r}+\mathbf{b}$. Therefore, letting $\mathbf{b}=-\mathbf{A} \mathbf{m}_{\mathbf{r}}$ shifts the mean to zero. Thus, we need

$$
\mathbf{b}=\left[\begin{array}{c}
\frac{-2}{\sqrt{3}} \\
0
\end{array}\right] .
$$

4. $(10+15=25$ points $)$
(a) This problem is essentially the same as the example I raised during class. Suppose $Z$ is Gaussian with zero mean and variance $A^{2}$. Let $N=Z+W$. We know $N$ is Gaussian with mean 0 and variance $\sigma^{2}+A^{2}$.
Thus, the system model now becomes

$$
Y=X+N .
$$

Then the posterior probability is given by

$$
\begin{aligned}
P[X=1 \mid Y=y] & =\frac{f_{Y \mid X}(y \mid X=1) P[X=1]}{f_{Y \mid X}(y \mid X=1) P[X=1]+f_{Y \mid X}(y \mid X=-1) P[X=-1]} \\
& =\frac{e^{-\frac{(y-1)^{2}}{2\left(A^{2}+\sigma^{2}\right)}}}{e^{-\frac{(y-1)^{2}}{2\left(A^{2}+\sigma^{2}\right)}}+e^{-\frac{(y+1)^{2}}{2\left(A^{2}+\sigma^{2}\right)}}} \\
& =\frac{e^{\frac{y}{A^{2}+\sigma^{2}}}}{e^{\frac{y}{A^{2}+\sigma^{2}}}+e^{\frac{-y}{A^{2}+\sigma^{2}}}}
\end{aligned}
$$

(b) Suppose now $Z$ is modeled as a binary discrete random variable equally likely to be $A$ or $-A$. This model reveals that the interference has a structure, rather than just Gaussian noise, that we can further exploit.
The posterior probability, by total probability theorem, is

$$
P[X=1 \mid Y=y]=\underbrace{P[X=1, Z=A \mid Y=y]}_{(1)}+\underbrace{P[X=1, Z=-A \mid Y=y]}_{(2)} .
$$

The first term is

$$
\begin{aligned}
P[X=1, Z=A \mid Y=y]= & f_{Y \mid X, Z}(y \mid X=1, Z=A) P[X=1, Z=A] / \\
& \left(\sum_{x \in\{1,-1\}, z \in\{A,-A\}} f_{Y \mid X, Z}(y \mid X=x, Z=z) P[X=x, Z=z]\right) \\
= & \frac{e^{-\frac{(y-1-A)^{2}}{2 \sigma^{2}}}}{e^{-\frac{(y-1-A)^{2}}{2 \sigma^{2}}}+e^{-\frac{(y-1+A)^{2}}{2 \sigma^{2}}}+e^{-\frac{(y+1-A)^{2}}{2 \sigma^{2}}}+e^{-\frac{(y+1+A)^{2}}{2 \sigma^{2}}}}
\end{aligned}
$$

We can similarly find term (2) as

$$
\begin{aligned}
P[X=1, Z=-A \mid Y=y]= & f_{Y \mid X, Z}(y \mid X=1, Z=-A) P[X=1, Z=-A] / \\
& \left(\sum_{x \in\{1,-1\}, z \in\{A,-A\}} f_{Y \mid X, Z}(y \mid X=x, Z=z) P[X=x, Z=z]\right) \\
= & \frac{e^{-\frac{(y-1+A)^{2}}{2 \sigma^{2}}}}{e^{-\frac{(y-1-A)^{2}}{2 \sigma^{2}}}+e^{-\frac{(y-1+A)^{2}}{2 \sigma^{2}}}+e^{-\frac{(y+1-A)^{2}}{2 \sigma^{2}}}+e^{-\frac{(y+1+A)^{2}}{2 \sigma^{2}}}} .
\end{aligned}
$$

It follows that

$$
P[X=1 \mid Y=y]=\frac{e^{-\frac{(y-1+A)^{2}}{2 \sigma^{2}}}+e^{-\frac{(y-1-A)^{2}}{2 \sigma^{2}}}}{e^{-\frac{(y-1-A)^{2}}{2 \sigma^{2}}}+e^{-\frac{(y-1+A)^{2}}{2 \sigma^{2}}}+e^{-\frac{(y+1-A)^{2}}{2 \sigma^{2}}}+e^{-\frac{(y+1+A)^{2}}{2 \sigma^{2}}}} .
$$

In part (a), the interference is modeled as Gaussian. And thus we treat it as noise. But on the other hand, in part (b), the interference is modeled more realistically, wherein the interference has a structure, rather than being treated as Gaussian noise, that we can further exploit to help us gain useful information about the desired signal $X$.
We will look deeper into this problem again when we talk about detection later in this course.

