**ECE 534 RANDOM PROCESSES** **FALL 2012**
**SOLUTIONS FOR PROBLEM SET 3**

## 1 (3.1) Rotation of a joint normal distribution yielding independence

(a) This part can be done just by writing the formula (3.8) for joint gaussian distribution and multiply the matrices:

$$f_X(x) = \frac{1}{(2\pi)^{\frac{m}{2}} |k|^{\frac{1}{2}}} \exp\left(-\frac{(x-\mu)^T K^{-1}(x-\mu)}{2}\right)$$

After substitution we have:

$$\begin{aligned} f_X(x) =& \frac{1}{(2\pi)} \exp\left(-\frac{\begin{bmatrix} x_1 - 10 & x_2 - 5 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 - 10 \\ x_2 - 5 \end{bmatrix}}{2}\right) \\ =& \frac{1}{(2\pi)} \exp\left(-\frac{x_1^2 + 2x_2^2 - 2x_1x_2 - 10x_1 + 50}{2}\right) \end{aligned}$$

(b) In order to find an appropriate linear transformation for these random variables which leads into new independant Gaussian tandom variables, we proceed like page 88 of lecture note by diagonalizing the covariance vecotor.

The eigenvalues of the covaoriance matrix can be calculated as $\frac{3\pm\sqrt{5}}{2}$ which corresponds to eigenvectors $U = \begin{bmatrix} 0.52 & -0.85 \\ -0.85 & -0.52 \end{bmatrix}$. If we consider $Y = U^T(X - \begin{bmatrix} 10 \\ 5 \end{bmatrix})$ then we have:

$$\begin{aligned} \mathbb{E}[YY^T] =& \mathbb{E}[U^T(X - \begin{bmatrix} 10 \\ 5 \end{bmatrix})(X - \begin{bmatrix} 10 \\ 5 \end{bmatrix})^T U] \\ =& U^T \mathrm{Cov}(X) U = \begin{bmatrix} \frac{3+\sqrt{5}}{2} & 0 \\ 0 & \frac{3-\sqrt{5}}{2} \end{bmatrix} \\ \mathbb{E}[Y] =& U^T \left(\mathbb{E}[X] - \begin{bmatrix} 10 \\ 5 \end{bmatrix}\right) = 0 \end{aligned}$$

## 2 (3.5) Conditional probabilities with joint Gaussian 1

(a) Since $(X, Y)$ have jointly Gaussian distribution, $P(X|Y)$ has also gaussian forms. Therefore, in order to find it's distribution, we only need to find its mean and variance. We have:

$$\begin{aligned}\mathbb{E}[X|Y] = \widehat{\mathbb{E}}[X|Y] =& \mathbb{E}[X] + \mathrm{Cov}(X,Y)\mathrm{Cov}^{-1}(Y,Y)\,(Y - \mathbb{E}[Y]) \\ =& 0 + \rho 1\,(y-0) = \rho y\end{aligned}$$

On the other hand:

$$\begin{aligned}\mathrm{Cov}(e) =& \mathrm{Cov}(X) - \mathrm{Cov}(X,Y)\mathrm{Cov}(Y,Y)^{-1}\mathrm{Cov}(Y,X) \\ =& 1-\rho^2\end{aligned}$$

Thereofore, $P(X|Y) \sim \mathcal{N}(\rho y, 1-\rho^2)$. This concludes that $P(X \leq 1|Y=y) = \Phi\left(\frac{1-\rho y}{\sqrt{1-\rho^2}}\right)$.

(b) For this part, we first expand the expression, then use calculations in part (a) especially the distribution of $P(X|Y)$:

$$\begin{aligned}\mathbb{E}[(X-Y)^2|Y] =& \mathbb{E}[X^2+Y^2-2XY|Y] = \mathbb{E}[X^2|Y] + Y^2 - 2Y\mathbb{E}[X|Y] \\ =& \mathbb{E}[X|Y]^2 + \mathrm{Var}[X|Y] + Y^2 - 2Y(\rho Y) \\ =& (\rho Y)^2 + \left(1-\rho^2\right) + Y^2 - 2Y(\rho Y) \\ =& \left(1-\rho^2\right) + (1-\rho)^2 Y^2\end{aligned}$$

## 3 Estimation of a symmetric binary signal in Gaussian noise

(a) We have $X \in \{+1,-1\}$ with probabilities $\frac{1}{2}$ and we observe $Y = X + \sqrt{\rho}Z$ where $Z \sim \mathcal{N}(0,1)$. To find $\mathbb{E}[X|Y]$, we have:

$$\begin{aligned}\mathbb{E}[X|Y] =& \int_{\mathbf{X}} f(x|Y)x\,dx \\ =& (+1)\mathbb{P}(X=1|Y) + (-1)\mathbb{P}(X=-1|Y) \\ =& \frac{\frac{1}{2}f_{Y|X}(y|1)}{\frac{1}{2}f_{Y|X}(y|1) + \frac{1}{2}f_{Y|X}(y-1)} - \frac{\frac{1}{2}f_{Y|X}(y|-1)}{\frac{1}{2}f_{Y|X}(y|1) + \frac{1}{2}f_{Y|X}(y|-1)} \\ =& \frac{f_{Y|X}(y|1) - f_{Y|X}(y|-1)}{f_{Y|X}(y|1) + f_{Y|X}(y-1)}.\end{aligned}$$

Since $X$ and $Z$ are independent, the distribution of $Y$ conditioned on $X = \pm 1$ is that of a Gaussian random variable with mean $\pm 1$ and variance $\rho$. Consequently,

$$\begin{aligned}\mathbb{E}[X|Y] =& \frac{\exp(-\frac{(Y-1)^2}{2\rho}) - \exp(-\frac{(Y+1)^2}{2\rho})}{\exp(-\frac{(Y-1)^2}{2\rho}) + \exp(-\frac{(Y+1)^2}{2\rho})} \\ =& \tanh\left(\frac{Y}{\rho}\right)\end{aligned}$$

(b) For this part, we need to use the general formula for linear estimation:

$$\widehat{\mathbb{E}}[X|Y] = \mathbb{E}[X] + \text{Cov}(X,Y)\text{Cov}^{-1}(Y,Y)\,(Y - \mathbb{E}[Y])$$

We have:

$$\begin{aligned}
&\mathbb{E}[X] = \frac{1}{2}(1) + \frac{1}{2}(-1) = 0 \\
&\text{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \mathbb{E}[X^2 + \sqrt{(}\rho)XZ] = 1 \\
&\text{Cov}^{-1}(Y,Y) = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 = \mathbb{E}[X^2 + \rho Z^2 + 2\sqrt{\rho}XZ] = 1 + \rho
\end{aligned}$$

Which results in:

$$\widehat{\mathbb{E}}[X|Y] = \frac{1}{1+\rho}\,(Y - \mathbb{E}[Y])$$

## 4 Estimation of a signal from multiple noisy measurements

(a) Since all variables have zero mean, the linear MMSE estimator of $X$ form $Y_1, \ldots, Y_n$ has the form

$$\widehat{\mathbb{E}}[X|Y] = \sum_{i=1}^{n} a_i Y_i.$$

A "clever" way to obtain the estimator is to exploit the symmetry in the problem: since each $Y_i$ has the same distribution, the coefficients $a_i$ must all be the same: $a_1 = \ldots = a_n = a$. By the orthogonality principle, we must have

$$\mathbb{E}\left[\left(X - a\sum_{i=1}^{n} Y_i\right) Y_j\right] = 0, \qquad \forall j$$

Using the fact that $Y_i = X + Z_i$ for each $i$ and the fact that the $Z_i$'s are independent with zero mean and variance $\sigma_Z^2$, we have

$$\mathbb{E}\left[\left((1-an)X - a\sum_{i=1}^{n} Z_i\right)(X + Z_j)\right] = (1-an)\sigma_X^2 - a\sigma_Z^2 = 0.$$

Solving for $a$, we get

$$a = \frac{\sigma_X^2}{n\sigma_X^2 + \sigma_Z^2}.$$

Therefore,

$$\widehat{\mathbb{E}}[X|Y] = \frac{\sigma_X^2}{n\sigma_X^2 + \sigma_Z^2}\sum_{i=1}^{n} Y_i.$$

The direct method is to use the formula for linear estimation:

$$\widehat{\mathbb{E}}[X|Y] = \mathbb{E}[X] + \text{Cov}(X,Y)\text{Cov}^{-1}(Y,Y)\,(Y - \mathbb{E}[Y])$$

We have:

$$\begin{aligned}
\text{Cov}(X,Y) = \mathbb{E}[XY^T] - \mathbb{E}[X]\mathbb{E}[Y] =& \mathbb{E}\left[X \begin{bmatrix} X+Z_1 & X+Z_2 & \cdots & X+Z_n \end{bmatrix}\right] \\
=& \begin{bmatrix} \sigma_X^2 & \sigma_X^2 & \cdots & \sigma_X^2 \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
\text{Cov}^{-1}(Y,Y) = \left(\mathbb{E}[YY^T] - \mathbb{E}[Y]\mathbb{E}[Y]^T\right)^{-1} =& \left(\mathbb{E}\left[\begin{bmatrix} X+Z_1 & X+Z_2 \cdots X+Z_n \end{bmatrix} \begin{bmatrix} X+Z_1 \\ X+Z_2 \\ \cdot \\ \cdot \\ \cdot \\ X+Z_n \end{bmatrix}\right]\right)^{-1} \\
=& \begin{bmatrix} \sigma_X^2+\sigma_Z^2 & \sigma_X^2 & \cdots & \sigma_X^2 \\ \sigma_X^2 & \sigma_X^2+\sigma_Z^2 & \cdots & \sigma_X^2 \\ & & \cdots & \\ & \cdots & & \\ \sigma_X^2 & \sigma_X^2 & \cdots & \sigma_X^2+\sigma_Z^2 \end{bmatrix}^{-1}
\end{aligned}$$

Putting everything together, we have:

$$\text{Cov}(X,Y)\text{Cov}^{-1}(Y,Y) = \begin{bmatrix} \sigma_X^2 & \sigma_X^2 & \cdots & \sigma_X^2 \end{bmatrix} \begin{bmatrix} \sigma_X^2+\sigma_Z^2 & \sigma_X^2 & \cdots & \sigma_X^2 \\ \sigma_X^2 & \sigma_X^2+\sigma_Z^2 & \cdots & \sigma_X^2 \\ & & \cdots & \\ & \cdots & & \\ \sigma_X^2 & \sigma_X^2 & \cdots & \sigma_X^2+\sigma_Z^2 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{\sigma_X^2}{\sigma_X^2+\sigma_Z^2} & \frac{\sigma_X^2}{n\sigma_X^2+\sigma_Z^2} & \cdots & \frac{\sigma_X^2}{n\sigma_X^2+\sigma_Z^2} \end{bmatrix}$$

Therefore,

$$\widehat{\mathbb{E}}[X|Y] = \frac{\sigma_X^2}{n\sigma_X^2+\sigma_Z^2}\sum_{i=1}^{n} Y_i$$

(b) To find the MSE, use equation (3.7) in lecture notes:

$$e = X - \widehat{\mathbb{E}}(X|Y_1, \cdots, Y_n)$$

From this,

$$\begin{aligned}
\text{Cov}(e) =& \text{Cov}(X) - \text{Cov}(X,Y)\text{Cov}(Y,Y)^{-1}\text{Cov}(Y,X) \\
=& \sigma_X^2 - \begin{bmatrix} \sigma_X^2 & \sigma_X^2 & \cdots & \sigma_X^2 \end{bmatrix} \begin{bmatrix} \sigma_X^2+\sigma_Z^2 & \sigma_X^2 & \cdots & \sigma_X^2 \\ \sigma_X^2 & \sigma_X^2+\sigma_Z^2 & \cdots & \sigma_X^2 \\ & & \cdots & \\ & \cdots & & \\ \sigma_X^2 & \sigma_X^2 & \cdots & \sigma_X^2+\sigma_Z^2 \end{bmatrix}^{-1} \begin{bmatrix} \sigma_X^2 \\ \sigma_X^2 \\ \cdot \\ \cdot \\ \cdot \\ \sigma_X^2 \end{bmatrix} \\
=& \frac{\sigma_X^2\sigma_Z^2}{n\sigma_X^2+\sigma_Z^2}
\end{aligned}$$

## 5 Data Processing and MMSE estmation

(a) We want to prove that if we have $Y = X + W_1$ and $Z = Y + W_2$, the $X \to Y \to Z$. We just neeed to rewrite the definiton here:

$$
\begin{aligned}
P(X=a, Z=c|Y=b) =& P(X=a, X+W_1+W_2=c|X+W_1=b) \\
=& \frac{P(X=a, X+W_1+W_2=c, X+W_1=b)}{P(X+W_1=b)} \\
=& \frac{P(X=a, W_1=b-X, W_2=c-X-W_1)}{P(X+W_1=b)} \\
=& \frac{P(X=a, W_1=b-a, W_2=c-b)}{P(X+W_1=b)} \\
=& \frac{P(X=a, W_1=b-a)P(W_2=c-b)}{P(X+W_1=b)} \\
=& \frac{P(X=a, X+W_1=b)P(b+W_2=c)}{P(X+W_1=b)} \frac{P(X+W_1=b)}{P(X+W_1=b)} \\
=& \frac{P(X=a, X+W_1=b)}{P(X+W_1=b)} \frac{P(b+W_2=c)P(X+W_1=b)}{P(X+W_1=b)} \\
=& P(X=a|X+W_1=b)\frac{P(b+W_2=c, X+W_1=b)}{P(X+W_1=b)} \\
=& P(X=a|X+W_1=b)P(b+W_2=c|X+W_1=b) \\
=& P(X=a|Y=b)P(X+W_1+W_2=c|X+W_1=b) \\
=& P(X=a|Y=b)P(Z=c|Y=b)
\end{aligned}
$$

(b) For this part:

$$
\begin{aligned}
\mathbb{E}[X|Y,Z] &= \int_{\mathbb{R}} x f_{X|Y,Z}(x|y,z)dx \\
&= \int_{\mathbb{R}} x \frac{f_{X,Y,Z}(x,y,z)}{f_{Y,Z}(y,z)}dx \\
&= \int_{\mathbb{R}} x \frac{f_{X,Z|Y}(x,z|y)f_Y(y)}{f_{Y,Z}(y,z)}dx \\
&= \int_{\mathbb{R}} x \frac{f_{X|Y}(x|y)f_{Z|Y}(z|y)f_Y(y)}{f_{Y,Z}(y,z)}dx \\
&= \int_{\mathbb{R}} x \frac{f_{X|Y}(x|y)f_{Y,Z}(y,z)}{f_{Y,Z}(y,z)}dx \\
&= \int_{\mathbb{R}} x f_{X|Y}(x|y)dx = \mathbb{E}[X|Y]
\end{aligned}
$$

(c) We want to prove that if $X \to Y \to Z$ then we have $\text{MMSE}(X|Z) - \text{MMSE}(X|Y) =$

$\mathbb{E}\left[(\mathbb{E}[X|Y]-\mathbb{E}[X|Z])^2\right]$. For this:

$$\begin{aligned}\mathrm{MMSE}(X|Z)-\mathrm{MMSE}(X|Y) =&\mathbb{E}\left[(X-\mathbb{E}[X|Z])^2\right]-\mathbb{E}\left[(X-\mathbb{E}[X|Y])^2\right]\\ =&\mathbb{E}\left[\mathbb{E}[X|Z]^2-\mathbb{E}[X|Y]^2+2X\mathbb{E}[X|Y]-2X\mathbb{E}[X|Z]\right]\\ =&\mathbb{E}\left[(\mathbb{E}[X|Z]^2-\mathbb{E}[X|Y])^2\right] \quad - \quad 2\mathbb{E}\left[(\mathbb{E}[X|Y]-X)(\mathbb{E}[X|Y]-\mathbb{E}[X|Z])\right]\end{aligned}$$

However, the second term on the right hand side can be proved to be equal to zero. We use the projection theorem to prove this:

$$\begin{aligned}\mathbb{E}\left[(\mathbb{E}[X|Y]-X)(\mathbb{E}[X|Y]-\mathbb{E}[X|Z])\right] =&\mathbb{E}\left[(\mathbb{E}[X|Y]-X)\mathbb{E}[X|Y]\right]+\mathbb{E}\left[\mathbb{E}[X|Z](X-\mathbb{E}[X|Y])\right]\\ =&\mathbb{E}\left[\mathbb{E}[X|Z](X-\mathbb{E}[X|Y])\right]=\mathbb{E}\left[\mathbb{E}[X|Z](X-\mathbb{E}[X|Y,Z])\right]=0\end{aligned}$$

(d) This can be easily proved from part (c) and the fact that $\mathbb{E}\left[(\mathbb{E}[X|Y]-\mathbb{E}[X|Z])^2\right]\geq 0$.

(e) The main idea here is that both the MMSE estimator $\mathbb{E}[X|Y]$ and the resulting MMSE $\mathbb{E}\left[(X-\mathbb{E}[X|Y])^2\right]$ depend only on the joint distribution of $X$ and $Y$. Suppose that $0<\rho_1<\rho_2$ and consider the random variables $X$, $Y_{\rho_1}$ and $Y_{\rho_2}$. We will construct another triple of random variables $\bar{X}$, $\bar{Y}_{\rho_1}$ and $\bar{Y}_{\rho_2}$, such that the following two properties hold:

1. $P_{XY_{\rho_1}}=P_{\bar{X}\bar{Y}_{\rho_1}}$ and $P_{XY_{\rho_2}}=P_{\bar{X}\bar{Y}_{\rho_2}}$
2. $\bar{X}\to\bar{Y}_{\rho_1}\to\bar{Y}_{\rho_2}$

By Property 1, $\mathrm{MMSE}(\bar{X}|\bar{Y}_{\rho_1})=\mathrm{MMSE}(X|Y_{\rho_1})$ and $\mathrm{MMSE}(\bar{X}|\bar{Y}_{\rho_2})=\mathrm{MMSE}(X|Y_{\rho_2})$. By Property 2 and the data processing inequality for MMSE estimation, $\mathrm{MMSE}(\bar{X}|\bar{Y}_{\rho_1})\geq\mathrm{MMSE}(\bar{X}|\bar{Y}_{\rho_2})$. Combining these two conclusions, we get the desired monotonicity property.

It remains to develop the desired construction. To that end, let $Z_1$ and $Z_2$ be two independent $N(0,1)$ random variables that are also independent of $X$. Then we let

$$\begin{aligned}\bar{X}&=X\\ \bar{Y}_{\rho_1}&=\bar{X}+\sqrt{\rho_1}Z_1\\ \bar{Y}_{\rho_2}&=\bar{X}+\sqrt{\rho_1}Z_1+\left(\sqrt{\rho_2-\rho_1}\right)Z_2.\end{aligned}$$

Then it is easy to see that $\bar{X}$ and $\bar{Y}_{\rho_1}$ have the same joint distribution as $X$ and $Y_{\rho_1}$. Also, the fact that $Z_1$ and $Z_2$ are i.i.d. $N(0,1)$ random variables implies that $\sqrt{\rho_1}Z_1+\sqrt{\rho_2-\rho_1}Z_2$ has the same distribution as $\sqrt{\rho_2}Z$. Consequently, $\bar{X}$ and $\bar{Y}_{\rho_2}$ have the same joint distribution as $X$ and $Y_{\rho_2}$. Thus, Property 1 is satisfied. Property 2 is satisfied by part (a), since by construction $\bar{Y}_{\rho_2}=\bar{Y}_{\rho_1}+W_2$ with $W_2=\sqrt{\rho_2-\rho_1}Z_1$, which is independent of $\bar{Y}_{\rho_1}$.