

### Problem 1

We know that independent random variables are uncorrelated. Argue that uncorrelated jointly Gaussian random variables are independent.

Hint: do this for two random variables first. For  $n$  random variables, you might find it easier to use the characteristic function.

**Solution:** Let  $\bar{U} = (U_1, \dots, U_n)^T$  be the  $n$  uncorrelated jointly Gaussian random variables. Let  $K$  be the covariance matrix of  $\bar{U}$  where for each  $i \in [n]$  we have  $Z_i \sim N(\mu_i, \sigma_i^2)$ . So  $\bar{U} = \bar{\mu} + \bar{Z}$  where  $\bar{Z} = (Z_1, \dots, Z_n)^T$  and  $\bar{Z}$  is zero mean Gaussian random variables. Since the Gaussian random variables are uncorrelated the matrix  $K$  is diagonal. Hence the  $K^{-1}$  is also diagonal. Then we know the density function of  $\bar{U}$  is

$$f_{\bar{U}}(\bar{u}) = \frac{\exp\left[-\frac{1}{2}(\bar{u} - \bar{\mu})^T K^{-1}(\bar{u} - \bar{\mu})\right]}{(2\pi)^{\frac{n}{2}} \sqrt{\det K}}$$

Since  $K$  is diagonal

$$K = \begin{bmatrix} \sigma_1^2 & & \\ & \sigma_2^2 & \\ & & \ddots \\ & & & \sigma_n^2 \end{bmatrix} \Rightarrow K^{-1} = \begin{bmatrix} \frac{1}{\sigma_1^2} & & \\ & \frac{1}{\sigma_2^2} & \\ & & \ddots \\ & & & \frac{1}{\sigma_n^2} \end{bmatrix}$$

Therefore we have

$$(\bar{u} - \bar{\mu})^T K^{-1}(\bar{u} - \bar{\mu}) = \sum_{i=1}^n (u_i - \mu_i) \frac{1}{\sigma_i^2} (u_i - \mu_i) = \sum_{i=1}^n \frac{(u_i - \mu_i)^2}{\sigma_i^2}$$

Hence we have

$$f_{\bar{U}}(\bar{u}) = \frac{\exp\left[-\frac{1}{2} \sum_{i=1}^n \frac{(u_i - \mu_i)^2}{\sigma_i^2}\right]}{(2\pi)^{\frac{n}{2}} \sqrt{\det K}} = \frac{\prod_{i=1}^n \exp\left[-\frac{1}{2} \frac{(u_i - \mu_i)^2}{\sigma_i^2}\right]}{(2\pi)^{\frac{n}{2}} \sqrt{\prod_{i=1}^n \sigma_i^2}} = \prod_{i=1}^n \frac{\exp\left[-\frac{(u_i - \mu_i)^2}{2\sigma_i^2}\right]}{\sqrt{2\pi\sigma_i^2}} = \prod_{i=1}^n f_{U_i}(u_i)$$

Therefore  $U_i$ 's are independent. ■

### Problem 2

- (i) \* Let  $X$  and  $Y$  be independent random variables.  $X_1 \sim N(0, 1)$ ; and  $Y = +1$  with probability  $p$  and  $Y = -1$  with probability  $1 - p$ . We define  $X_2 = YX_1$ . Is  $X_2$  Gaussian? Are  $X_1, X_2$  jointly Gaussian? Justify your answers.

[See Example 3.3.4 from [G] for a solution]

- (ii) Repeat (i) if  $X_1 \sim N(m, 1)$  and  $m > 0$

**Solution:** We know for any random variable  $Z \sim N(\mu, \sigma^2)$  the characteristic function of  $Z$  is  $\mathbb{E}[\exp(itZ)] = \exp(it\mu - \frac{1}{2}\sigma^2 t^2)$ .

Now we know  $X_1 \sim N(m, 1)$  where  $m > 0$ . Therefore  $\mathbb{E}[X_1] = m$  and  $\text{Var}[X_1] = 1$ . Therefore  $\mathbb{E}[X_1^2] = \text{Var}[X_1] + \mathbb{E}[X_1]^2 = 1 + m^2$ . Also for  $Y$  we have  $\mathbb{E}[Y] = p - (1 - p) = 2p - 1$  and  $\mathbb{E}[Y^2] = p + (1 - p) = 1$ . Now we will calculate the mean and the variance and the characteristic function of  $X_2 = X_1 Y$ .

$$\mathbb{E}[X_2] = \mathbb{E}[X_1 Y] = \mathbb{E}[X_1] \mathbb{E}[Y] = (2p - 1)m$$

Now

$$\mathbb{E}[X_2^2] = \mathbb{E}[X_1^2 Y^2] = \mathbb{E}[X_1^2] \mathbb{E}[Y^2] = m^2 + 1$$

Hence we have

$$\text{Var}[X_2] = \mathbb{E}[X_2^2] - \mathbb{E}[X_2]^2 = m^2 + 1 - (2p - 1)^2 m^2 = m^2 + 1 - (4p^2 - 4p + 1)m^2 = 1 - 4m^2(p^2 - p)$$

Hence if  $X_2$  is Gaussian then we have  $X_2 \sim N((2p - 1)m, 1 - 4m^2(p^2 - p))$ . Then the characteristic function of  $X_2$  would have become  $\exp\left(it(2p - 1)m - \frac{t^2}{2}(1 - 4m^2(p^2 - p))\right)$ . Now let's calculate the characteristic function of  $X_2$ .

$$\mathbb{E}[\exp(itX_2)] = \mathbb{E}[\exp(itX_1)] \mathbb{E}[\exp(itY)] = \exp\left(itm - \frac{t^2}{2}\right) [pe^{it} + (1 - p)e^{-it}]$$

So comparing the two equations we have

$$\begin{aligned} \exp\left(it(2p - 1)m - \frac{t^2}{2}(1 - 4m^2(p^2 - p))\right) &= \exp\left(itm - \frac{t^2}{2}\right) [pe^{it} + (1 - p)e^{-it}] \\ \implies \exp\left(it(2p - 1)m - \frac{t^2}{2}(1 - 4m^2(p^2 - p)) - \left[itm - \frac{t^2}{2}\right]\right) &= pe^{it} + (1 - p)e^{-it} \\ \implies \exp\left(2it(p - 1)m + \frac{t^2}{2}(4m^2(p^2 - p))\right) &= pe^{it} + (1 - p)e^{-it} \\ \implies \exp(2it(p - 1)m + 2t^2 m^2(p^2 - p)) &= pe^{it} + (1 - p)e^{-it} \end{aligned}$$

Now notice that  $p \leq 1$ . Hence  $p - 1 \leq 0$  and  $p^2 - p \leq 0$ . Therefore we have

$$2it(p - 1)m + 2t^2 m^2(p^2 - p) \leq 0 \implies \exp(2it(p - 1)m + 2t^2 m^2(p^2 - p)) \leq 1$$

But in the *RHS* we have

$$pe^{it} + (1 - p)e^{-it} = p(\cos t + i \sin t) + (1 - p)(\cos t - i \sin t) = \cos t + i(2p - 1) \sin t$$

Therefore  $|pe^{it} + (1 - p)e^{-it}| = \sqrt{1 + (2p - 1)^2} > 1$ . But this is not possible. Hence contradiction.  $X_2$  is not Gaussian.

If  $X_1, X_2$  is jointly Gaussian then the marginal distribution on  $X_2$  is also Gaussian. Since we know the marginal distribution on  $X_2$  is not Gaussian we have  $X_1, X_2$  are not jointly Gaussian. ■

### Problem 3 [G] Exercise 3.8

- (a) Let  $[K] = \begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix}$ . Show that 1 and  $\frac{1}{2}$  are eigenvalues of  $[K]$  and find the normalized eigenvectors. Express  $[K]$  as  $[Q]\Lambda Q^{-1}$ , where  $[\Lambda]$  is diagonal and  $[Q]$  is orthonormal.
- (b) Let  $[K'] = \alpha[K]$  for real  $\alpha \neq 0$ . Find the eigenvalues and eigenvectors of  $[K']$ . Don't not use brute force - think!
- (c) Find the eigenvalues and eigenvectors of  $[K^m]$ , where  $[K^m]$  is the  $m$ th power of  $[K]$ .

**Solution:**

- (a) Let the vector  $\begin{bmatrix} 1 & 1 \end{bmatrix}^T$  and  $\begin{bmatrix} 1 & -1 \end{bmatrix}^T$ . We claim they are the eigenvectors corresponding to eigenvalues 1 and  $-1$  respectively.

$$\begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.75 + 0.25 \\ 0.25 + 0.75 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0.75 - 0.25 \\ 0.25 - 0.75 \end{bmatrix} = \begin{bmatrix} 0.5 \\ -0.5 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Hence  $\begin{bmatrix} 1 & 1 \end{bmatrix}^T$  and  $\begin{bmatrix} 1 & -1 \end{bmatrix}^T$  are indeed eigenvector corresponding to eigenvalues 1 and  $-1$  respectively.

Now the vectors  $\begin{bmatrix} 1 & 1 \end{bmatrix}^T$  and  $\begin{bmatrix} 1 & -1 \end{bmatrix}^T$  are orthogonal but they are not normalized vectors. Hence consider the vectors  $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \end{bmatrix}^T$  and  $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \end{bmatrix}^T$ . They are orthogonal and also they are normalized. Hence they are orthonormal. Hence we claim  $[Q] = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ . Since we already knew the eigenvalues we also have  $[\Lambda] = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$ . First we will show that  $[Q^T] = [Q^{-1}]$ . Now  $\det[Q] = \left(\frac{1}{\sqrt{2}}\right)^2 (1 \times (-1) - 1 \times 1) = \frac{1}{2} \times (-2) = -1$ .  
Now

$$[Q^{-1}] = \frac{1}{\det[Q]} \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = [Q^T]$$

So now

$$[Q\Lambda Q^{-1}] = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0.5 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 0.5 \\ 1 & -0.5 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1.5 & 0.5 \\ 0.5 & 1.5 \end{bmatrix} = \begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix} = [K]$$

(b) If  $v$  is an eigenvector with corresponding eigenvalue  $\lambda$  of  $[K]$  then we have

$$[K']v = \alpha[K]v = \alpha\lambda v = (\alpha\lambda)v$$

So  $v$  is also an eigenvector of  $[K']$  but the corresponding eigenvalue is  $\alpha\lambda$ . Since by the previous part we know the eigenvector of  $[K]$  are  $\begin{bmatrix} 1 & 1 \end{bmatrix}^T$  and  $\begin{bmatrix} 1 & -1 \end{bmatrix}^T$  with corresponding eigenvalues 1 and  $\frac{1}{2}$  respectively the eigenvectors of  $[K']$  are the same  $\begin{bmatrix} 1 & 1 \end{bmatrix}^T$  and  $\begin{bmatrix} 1 & -1 \end{bmatrix}^T$  with corresponding eigenvalues  $\alpha$  and  $\frac{\alpha}{2}$  respectively.

(c) If  $v$  is an eigenvector with corresponding eigenvalue  $\lambda$  of  $[K]$  then we have

$$[K^m]v = [K^{m-1}][K]v = [K^{m-1}]\alpha v = \alpha[K^{m-1}]v = \alpha^2[K^{m-2}]v = \dots = \alpha^{m-1}[K]v = \alpha^m v$$

Therefore  $v$  is also an eigenvector of  $[K^m]$  but the corresponding eigenvalue is  $\lambda^m$ . Since by the part (a) we know the eigenvector of  $[K]$  are  $\begin{bmatrix} 1 & 1 \end{bmatrix}^T$  and  $\begin{bmatrix} 1 & -1 \end{bmatrix}^T$  with corresponding eigenvalues 1 and  $-1$  respectively the eigenvectors of  $[K^m]$  are the same  $\begin{bmatrix} 1 & 1 \end{bmatrix}^T$  and  $\begin{bmatrix} 1 & -1 \end{bmatrix}^T$  with corresponding eigenvalues 1 and  $\frac{1}{2^m}$  respectively.

■

#### Problem 4

We derived the p.d.f. of a jointly Gaussian random vector  $X = AW$ , where  $A$  is an  $n \times n$  matrix. We used the fact  $A$  is invertible. How would you precisely describe the distribution of  $X$  if  $A$  is not invertible? Describe the underlying geometry of the distribution of  $X$ . Use the following  $A$  as an example:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 2 & 3 & 4 \end{pmatrix}$$

**Solution:**

■

#### Problem 5 [G] Problem 3.9

Let  $X$  and  $Y$  be jointly Gaussian with means  $m_X, m_Y$ , variances  $\sigma_X^2, \sigma_Y^2$ , and normalized covariance  $\rho$ . Find the conditional density  $f_{X|Y}(x | y)$ .

**Solution:** We have  $\mathbb{E}[X] = m_X$  and  $\mathbb{E}[Y] = m_Y$ . Hence  $\rho = \frac{\mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]}{\sigma_X \sigma_Y} = \frac{\mathbb{E}[(X - m_X)(Y - m_Y)]}{\sigma_X \sigma_Y}$ . So  $\text{Cov}(X, Y) = \rho \sigma_X \sigma_Y$ . Hence the covariance matrix is

$$K = \begin{bmatrix} \sigma_X^2 & \rho \sigma_X \sigma_Y \\ \rho \sigma_X \sigma_Y & \sigma_Y^2 \end{bmatrix}$$

Now  $\det K = \sigma_X^2 \sigma_Y^2 - \rho^2 \sigma_X^2 \sigma_Y^2 = \sigma_X^2 \sigma_Y^2 (1 - \rho^2)$ . Then

$$K^{-1} = \frac{1}{\sigma_X^2 \sigma_Y^2 (1 - \rho^2)} \begin{bmatrix} \sigma_Y^2 & -\rho \sigma_X \sigma_Y \\ -\rho \sigma_X \sigma_Y & \sigma_X^2 \end{bmatrix} = \frac{1}{1 - \rho^2} \begin{bmatrix} \frac{1}{\sigma_X^2} & -\frac{\rho}{\sigma_X \sigma_Y} \\ -\frac{\rho}{\sigma_X \sigma_Y} & \frac{1}{\sigma_Y^2} \end{bmatrix}$$

Now we know the joint density function of  $X, Y$  is

$$\begin{aligned} f_{X,Y}(x, y) &= \frac{1}{2\pi \sqrt{\det K}} \exp \left( -\frac{1}{2(1 - \rho^2)} \begin{bmatrix} x - m_X & y - m_Y \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma_X^2} & -\frac{\rho}{\sigma_X \sigma_Y} \\ -\frac{\rho}{\sigma_X \sigma_Y} & \frac{1}{\sigma_Y^2} \end{bmatrix} \begin{bmatrix} x - m_X \\ y - m_Y \end{bmatrix} \right) \\ &= \frac{1}{2\pi \sigma_X \sigma_Y \sqrt{1 - \rho^2}} \exp \left( -\frac{1}{2(1 - \rho^2)} \begin{bmatrix} x - m_X & y - m_Y \end{bmatrix} \begin{bmatrix} \frac{x - m_X}{\sigma_X^2} - \rho \frac{y - m_Y}{\sigma_X \sigma_Y} \\ -\rho \frac{x - m_X}{\sigma_X \sigma_Y} + \frac{y - m_Y}{\sigma_Y^2} \end{bmatrix} \right) \\ &= \frac{1}{2\pi \sigma_X \sigma_Y \sqrt{1 - \rho^2}} \exp \left( -\frac{(x - m_X) \left[ \frac{x - m_X}{\sigma_X^2} - \rho \frac{y - m_Y}{\sigma_X \sigma_Y} \right] + (y - m_Y) \left[ -\rho \frac{x - m_X}{\sigma_X \sigma_Y} + \frac{y - m_Y}{\sigma_Y^2} \right]}{2(1 - \rho^2)} \right) \\ &= \frac{1}{2\pi \sigma_X \sigma_Y \sqrt{1 - \rho^2}} \exp \left( -\frac{\frac{(x - m_X)^2}{\sigma_X^2} - \rho \frac{(x - m_X)(y - m_Y)}{\sigma_X \sigma_Y} - \rho \frac{(x - m_X)(y - m_Y)}{\sigma_X \sigma_Y} + \frac{(y - m_Y)^2}{\sigma_Y^2}}{2(1 - \rho^2)} \right) \\ &= \frac{1}{2\pi \sigma_X \sigma_Y \sqrt{1 - \rho^2}} \exp \left( -\frac{1}{2(1 - \rho^2)} \left[ \frac{(x - m_X)^2}{\sigma_X^2} - 2\rho \frac{(x - m_X)(y - m_Y)}{\sigma_X \sigma_Y} + \frac{(y - m_Y)^2}{\sigma_Y^2} \right] \right) \end{aligned}$$

Now we have  $f_Y(y) = \frac{1}{\sigma_Y \sqrt{2\pi}} \exp \left( -\frac{(y - m_Y)^2}{2\sigma_Y^2} \right)$ . We know for conditional density function  $f_{X|Y}(x | y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$ .

Hence we have

$$\begin{aligned}
f_{X|Y}(x|y) &= \frac{\frac{1}{2\pi\sigma_x\sigma_Y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)} \left[ \frac{(x-m_X)^2}{\sigma_x^2} - 2\rho \frac{(x-m_X)(y-m_Y)}{\sigma_x\sigma_Y} + \frac{(y-m_Y)^2}{\sigma_Y^2} \right] \right)}{\frac{1}{\sigma_Y\sqrt{2\pi}} \exp\left(-\frac{(y-m_Y)^2}{2\sigma_Y^2}\right)} \\
&= \frac{1}{\sigma_x\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{\left[ \frac{(x-m_X)^2}{\sigma_x^2} - 2\rho \frac{(x-m_X)(y-m_Y)}{\sigma_x\sigma_Y} + \frac{(y-m_Y)^2}{\sigma_Y^2} - (1-\rho^2) \frac{(y-m_Y)^2}{2\sigma_Y^2} \right]}{2(1-\rho^2)}\right) \\
&= \frac{1}{\sigma_x\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{\left[ \frac{(x-m_X)^2}{\sigma_x^2} - 2\rho \frac{(x-m_X)(y-m_Y)}{\sigma_x\sigma_Y} + \rho^2 \frac{(y-m_Y)^2}{\sigma_Y^2} \right]}{2(1-\rho^2)}\right) \\
&= \frac{1}{\sigma_x\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{1}{2(1-\rho^2)} \left[ \frac{x-m_X}{\sigma_x} - \rho \frac{y-m_Y}{\sigma_Y} \right]^2\right) \\
&= \frac{1}{\sigma_x\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{1}{2\sigma_X^2(1-\rho^2)} \left[ x - \left( \rho \frac{\sigma_X}{\sigma_Y} (y-m_Y) + m_X \right) \right]^2\right)
\end{aligned}$$

Hence we have  $X | Y = y \sim N\left(\rho \frac{\sigma_X}{\sigma_Y} (y - m_Y) + m_X, \sigma_X^2 (1 - \rho^2)\right)$ . ■

In the next two problems we will use a common model for communication systems. The transmitted signal  $\vec{X}$  is a Gaussian random vector of size  $m$  (vector since there are several, say  $m$ , transmit antennas and each component of the vector stands for the input to a separate antenna). The signal goes over a linear and additive Gaussian noise channel and is picked up by a receiver which also has  $n$  antennas. The received vector of length  $n$  has the form.

$$\vec{Y} = H\vec{X} + \vec{Z}, \quad (1)$$

where  $H$  is a constant  $n \times m$  vector and  $\vec{Z}$  is a Gaussian random vector of size  $n$  and independent of  $\vec{X}$ .

### Problem 6

Let us first consider the simpler case of  $m = 1$  and  $n = 2$ . So  $X$  is a scalar random variable. Let  $X$  have the standard normal distribution  $N(0, 1)$ . The received signals are

$$Y_i = h_i X + Z_i, \quad i = 1, 2,$$

where  $Z_i \sim N(0, \sigma^2)$  are i.i.d and independent of  $X$ . And  $h_i$ 's are constants which represent the channel "gains" from the transmit antenna to the receive antennas.

- Find the conditional joint distribution of  $Y_1, Y_2$  conditioned on  $X = x$ .
- Find the conditional joint distribution of  $X$  conditioned on  $Y_1 = y_1, Y_2 = y_2$ .
- Using (b), what is your estimate of the transmitted signal  $X$  if you are told that the receive antennas observed  $Y_1 = y_1, Y_2 = y_2$ . **Interpret your results.** Does your answer make intuitive sense? What happens to the estimate when the noise variance  $\sigma^2$  becomes small? or large?

**Solution:** Now  $\tilde{Z} = [X \ Z_1 \ Z_2]^T$  forms independent zero mean Gaussian 3-random vectors since  $X \sim N(0, 1)$ ,  $Z_1 \sim N(0, \sigma^2)$ ,  $Z_2 \sim N(0, \sigma^2)$ . Hence the covariance matrix of  $\tilde{Z}$  is

$$\mathbb{E}[\tilde{Z}\tilde{Z}^T] = \begin{bmatrix} 1 & & \\ & \sigma^2 & \\ & & \sigma^2 \end{bmatrix}$$

Consider the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 \\ h_1 & 1 & 0 \\ h_2 & 0 & 1 \end{bmatrix}$$

Then we have

$$\begin{bmatrix} X \\ Y_1 \\ Y_2 \end{bmatrix} = A \begin{bmatrix} X \\ Z_1 \\ Z_2 \end{bmatrix}$$

Hence the 3-random vector  $\tilde{Y} = [X \ Y_1 \ Y_2]^T$  is a zero mean Gaussian 3-random vectors. Now let  $K$  denote the covariance matrix of  $\tilde{Y}$ . Then

$$K = \mathbb{E}[\tilde{Y}\tilde{Y}^T] = \mathbb{E}[A\tilde{Z}\tilde{Z}^T A^T] = A\mathbb{E}[\tilde{Z}\tilde{Z}^T]A^T = \begin{bmatrix} 1 & 0 & 0 \\ h_1 & 1 & 0 \\ h_2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{bmatrix} \begin{bmatrix} 1 & h_1 & h_2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & h_1 & h_2 \\ h_1 & h_1 + \sigma^2 & h_1 h_2 \\ h_2 & h_1 h_2 & h_2^2 + \sigma^2 \end{bmatrix}$$

Now

$$K = \begin{bmatrix} K_X & K_{X \cdot Y} \\ K_{X \cdot Y}^T & K_Y \end{bmatrix} = \begin{bmatrix} 1 & h_1 & h_2 \\ h_1 & h_1 + \sigma^2 & h_1 h_2 \\ h_2 & h_1 h_2 & h_2^2 + \sigma^2 \end{bmatrix}$$

Therefore  $K_X = [1]$ ,  $K_Y = \begin{bmatrix} h_1 + \sigma^2 & h_1 h_2 \\ h_1 h_2 & h_2^2 + \sigma^2 \end{bmatrix}$  and  $K_{X \cdot Y} = K_{Y \cdot X}^T = [h_1 \ h_2]$ .

- (a) Let  $\bar{Y} = [Y_1 \ Y_2]^T$ . Then we are asked to find  $\bar{Y} \mid X = x$ . We know  $\bar{Y} \mid X = x$  is Gaussian bivariate random vector. The mean of  $\bar{Y} \mid X = x$  is

$$K_{Y \cdot X} K_X^{-1} x = K_{X \cdot Y}^T K_X^{-1} x = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} x = \begin{bmatrix} h_1 x \\ h_2 x \end{bmatrix}$$

The variance of  $\bar{Y} \mid X = x$  is

$$K_Y - K_{Y \cdot X} K_X^{-1} K_{Y \cdot X}^T = \begin{bmatrix} h_1 + \sigma^2 & h_1 h_2 \\ h_1 h_2 & h_2^2 + \sigma^2 \end{bmatrix} - \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \begin{bmatrix} h_1 & h_2 \end{bmatrix} = \begin{bmatrix} h_1 + \sigma^2 & h_1 h_2 \\ h_1 h_2 & h_2^2 + \sigma^2 \end{bmatrix} - \begin{bmatrix} h_1^2 h_1 h_2 & h_1^2 h_2 \\ h_1 h_2 & h_2^2 \end{bmatrix} = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}$$

Therefore we have  $Y_1 \mid X = x \sim N(h_1 x, \sigma^2)$  and  $Y_2 \mid X = x \sim N(h_2 x, \sigma^2)$ .

- (b) Let  $\bar{y} = [y_1 \ y_2]^T$ . We are asked to find  $X \mid \bar{Y} = \bar{y}$ . We know  $X \mid \bar{Y} = \bar{y}$  is a Gaussian distribution. But we will find the mean and the variance of the distribution now. First we will find  $K_Y^{-1}$ .

$$K_Y^{-1} = \begin{bmatrix} h_1 + \sigma^2 & h_1 h_2 \\ h_1 h_2 & h_2^2 + \sigma^2 \end{bmatrix}^{-1} = \frac{1}{\sigma^4 + \sigma^2(h_1^2 + h_2^2)} \begin{bmatrix} h_2^2 + \sigma^2 & -h_1 h_2 \\ -h_1 h_2 & h_1^2 + \sigma^2 \end{bmatrix}$$

The mean of  $X \mid \bar{Y} = \bar{y}$  is

$$K_{X \cdot Y} K_Y^{-1} \bar{y} = \frac{1}{\sigma^4 + \sigma^2(h_1^2 + h_2^2)} [h_1 \ h_2] \begin{bmatrix} h_2^2 + \sigma^2 & -h_1 h_2 \\ -h_1 h_2 & h_1^2 + \sigma^2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \frac{\sigma^2(h_1 y_1 + h_2 y_2)}{\sigma^4 + \sigma^2(h_1^2 + h_2^2)} = \frac{h_1 y_1 + h_2 y_2}{\sigma^2 + h_1^2 + h_2^2}$$

The variance of  $X \mid \bar{Y} = \bar{y}$  is

$$K_X - K_{X \cdot Y} K_Y^{-1} K_{X \cdot Y}^T = 1 - \frac{1}{\sigma^4 + \sigma^2(h_1^2 + h_2^2)} [h_1 \ h_2] \begin{bmatrix} h_2^2 + \sigma^2 & -h_1 h_2 \\ -h_1 h_2 & h_1^2 + \sigma^2 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = 1 - \frac{\sigma^2(h_1 + h_2)}{\sigma^4 + \sigma^2(h_1^2 + h_2^2)} = \frac{\sigma^2}{\sigma^2 + h_1^2 + h_2^2}$$

Therefore we have  $X \mid \bar{Y} = \bar{y} \sim N\left(\frac{h_1 y_1 + h_2 y_2}{\sigma^2 + h_1^2 + h_2^2}, \frac{\sigma^2}{\sigma^2 + h_1^2 + h_2^2}\right)$

(c) Hence the estimated transmitted signal  $X$  if observed  $Y_1 = y_1$  and  $Y_2 = y_2$  is  $\frac{h_1 y_1 + h_2 y_2}{\sigma^2 + h_1^2 + h_2^2}$ .

Now

$$\lim_{\sigma^2 \rightarrow 0} \frac{h_1 y_1 + h_2 y_2}{\sigma^2 + h_1^2 + h_2^2} = \frac{h_1 y_1 + h_2 y_2}{h_1^2 + h_2^2}$$

Hence if  $\sigma^2$  becomes very small then the estimated transmitted signal is  $\frac{h_1 y_1 + h_2 y_2}{h_1^2 + h_2^2}$ .

If  $\sigma^2$  becomes large

$$\lim_{\sigma^2 \rightarrow \infty} \frac{h_1 y_1 + h_2 y_2}{\sigma^2 + h_1^2 + h_2^2} = 0$$

then the estimated transmitted signal is 0.

■

### Problem 7

Now consider the general model in (1) for general  $n, m$ . Let  $\vec{X} \sim N(\vec{0}, K_X)$ ,  $\vec{Z} \sim N(\vec{0}, K_Z)$  and  $\vec{Z}$  is independent of  $\vec{X}$ .

- (a) Show that  $\vec{U} = (\vec{X}, \vec{Y})$  is jointly Gaussian. You may use any of the equivalent definitions we saw in class
- (b) Find a simple condition on  $H, K_X, K_Z$  so that  $K_U$  is invertible.
- (c) What is the conditional distribution of the input  $\vec{X}$  given the output  $\vec{Y} = \vec{y}$ .

**Solution:**

- (a) Now  $\hat{Z} = [\vec{X}^T, \vec{Z}^T]^T$  forms independent zero mean Gaussian  $(n+m)$ -random variable since  $\vec{X} \sim N(\vec{0}, K_X)$  and  $\vec{Z} \sim N(\vec{0}, K_Z)$ . Also denote  $\hat{Y} = [\vec{X}^T, \vec{Y}^T]^T$ . Now we know

$$\vec{Y} = H\vec{X} + \vec{Z} \implies \vec{Y} = [H \mid I_n] \begin{bmatrix} \vec{X} \\ \vec{Z} \end{bmatrix} \implies \begin{bmatrix} \vec{X} \\ \vec{Y} \end{bmatrix} = \underbrace{\begin{bmatrix} I_m & | & \\ H & | & I_n \end{bmatrix}}_A \begin{bmatrix} \vec{X} \\ \vec{Z} \end{bmatrix} \implies \hat{Y} = A\hat{Z}$$

Since  $\hat{Z}$  is zero mean Gaussian  $(n+m)$ -random vector  $\hat{Y}$  is also a zero mean Gaussian  $(n+m)$ -random vector. Hence  $\vec{U} = (\vec{X}, \vec{Y})$  is jointly Gaussian.

- (b) Now covariance matrix of  $\vec{U}$  or  $\hat{Y}$  is  $K_U$ . The covariance matrix of  $\hat{Z}$  is

$$\mathbb{E}[\hat{Z}\hat{Z}^T] = \begin{bmatrix} K_X & \\ & K_Z \end{bmatrix}$$

Then we have

$$K_U = \mathbb{E}[\hat{Y}\hat{Y}^T] = \mathbb{E}[A\hat{Z}\hat{Z}^T A^T] = A\mathbb{E}[\hat{Z}\hat{Z}^T]A^T = \begin{bmatrix} I_m & \\ H & I_n \end{bmatrix} \begin{bmatrix} K_X & \\ & K_Z \end{bmatrix} \begin{bmatrix} I_m & H^T \\ & I_n \end{bmatrix} = \begin{bmatrix} K_X & K_X H^T \\ HK_X & HK_X H^T + K_Z \end{bmatrix}$$

Let the inverse of  $K_U$  is

$$K_U^{-1} = \begin{bmatrix} P & Q \\ R & S \end{bmatrix} \implies K_U K_U^{-1} = \begin{bmatrix} K_X P + K_X H^T R & K_X Q + K_X H^T S \\ HK_X P + (HK_X H^T + K_Z)R & HK_X Q + (HK_X H^T + K_Z)S \end{bmatrix} = \begin{bmatrix} I_m & \\ & I_n \end{bmatrix}$$

Then we have

$$K_X P + K_X H^T R = I_m \implies K_X (P + H^T R) = I_m \implies K_X \text{ is invertible}$$

Now we have

$$HK_X P + (HK_X H^T + K_Z)R = 0 \implies HK_X (P + H^T R) + K_Z R = 0 \implies H + K_Z R = 0$$

We also have

$$HK_X Q + (HK_X H^T + K_Z)S = I_n \implies H(K_X Q + K_X H^T S) + K_Z S = I_n \implies K_Z S = I_n \implies K_Z \text{ is invertible}$$

If  $K_X, K_Z$  are invertible then we have  $S = K_Z^{-1} \cdot H + K_Z R = 0 \implies R = -K_Z^{-1} H$ .

$$K_X Q + K_X H^T K_Z^{-1} = 0 \implies Q = -H^T K_Z^{-1}$$

And finally

$$P + H^T R = K_X^{-1} \implies P = K_X^{-1} + H^T K_Z^{-1} H$$

Therefore if  $K_X, K_Y$  and  $HK_X H^T + K_Z$  are invertible then  $K_U$  becomes invertible.

(c) We have  $K_U = \begin{bmatrix} K_X & K_X H^T \\ HK_X & HK_X H^T + K_Z \end{bmatrix}$ . Also from this we get

$$K_U^{-1} = \begin{bmatrix} K_X^{-1} + H^T K_Z^{-1} H & -H^T K_Z^{-1} \\ -K_Z^{-1} H & K_Z^{-1} \end{bmatrix}$$

Now we know  $\vec{X} \mid \vec{Y} = \vec{y}$  is a Gaussian  $m$ -random variable. The mean of  $\vec{X} \mid \vec{Y} = \vec{y}$  is  $P^{-1}Q = -(K_X^{-1} + H^T K_Z^{-1} H)^{-1} H^T K_Z^{-1}$ . And the variance is  $(K_X^{-1} + H^T K_Z^{-1} H)^{-1}$ . Therefore we have the distribution function of  $\vec{X} \mid \vec{Y} = \vec{y}$  is  $N(-(K_X^{-1} + H^T K_Z^{-1} H)^{-1} H^T K_Z^{-1}, (K_X^{-1} + H^T K_Z^{-1} H)^{-1})$ .

■