Homework 7

Due: Thursday, November 30, 2023, 1:00pm on Gradescope

Please upload your answers timely to Gradescope. Start a new page for every problem. For the programming/simulation questions you can use any reasonable programming language. Comment your source code and include the code and a brief overall explanation with your answers.

1. Exercises 3.12 from the text. (12 points)

   Let $\mathbf{X}$ and $\mathbf{Y}$ be jointly-Gaussian rv.s with means $\mathbf{m_X}$ and $\mathbf{m_Y}$, covariance matrices $K_\mathbf{X}$ and $K_\mathbf{Y}$ and cross covariance matrix $K_{\mathbf{X}\cdot\mathbf{Y}}$. Find the conditional probability density $f_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y})$. Assume that the covariance of $\begin{bmatrix}\mathbf{X}\\ \mathbf{Y}\end{bmatrix}$ is non-singular. Book Hint: Think of the fluctuations of $\mathbf{X}$ and $\mathbf{Y}$.

   Our Hint: You can first assume that $\mathbf{X}$ and $\mathbf{Y}$ have zero means to solve the problem, and then add in the means later. You may also assume that the inverse of the covariance matrix of $\begin{bmatrix}\mathbf{X}\\ \mathbf{Y}\end{bmatrix}$ is of the form

   $$K^{-1} = \begin{bmatrix} B & C \\ C^\top & D \end{bmatrix}$$

   where B is a square matrix of the same dimension as $\mathbf{X}$ and D is a square matrix of the same dimension as $\mathbf{Y}$. Compute your answer in terms of B, C and D.

   The cross covariance matrix of $\mathbf{X}$ and $\mathbf{Y}$ is defined as $K_{\mathbf{X}\cdot\mathbf{Y}} = \mathbb{E}[(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{Y} - \mathbb{E}\mathbf{Y}])^T]$ so the covariance matrix of the random vector $\begin{bmatrix}\mathbf{X}\\ \mathbf{Y}\end{bmatrix}$ is

   $$K = \begin{bmatrix} K_X & K_{X\cdot Y} \\ K_{X\cdot Y}^\intercal & K_Y \end{bmatrix}.$$

   Extra credit: You can even compute B, C and D in terms of the given covariance matrices.

2. Communication over colored Gaussian noise channel (15 points)

You want to communicate one bit over a Gaussian channel. You transmit $X = \sqrt{P}$ when $B = 1$ and $X = -\sqrt{P}$ when $B = 0$, where B is Bernoulli(1/2). The received signal at the receiver is $Y = X + Z$ where $Z \sim \mathcal{N}(0, N)$.

a. (2 points) What is the optimal decoding rule $\hat{B}(Y)$ at the receiver that minimizes $\mathbb{P}(B \neq \hat{B}(Y))$? What is the probability of error of this optimal decoder?

Realizing that the channel is not white, i.e. the noise samples over consecutive channel uses are correlated, you decide to measure the channel noise in the next time slot and use it to improve your probability of error. Now you have two observations $Y_1 = X + Z_1$ and $Y_2 = Z_2$ where Z_1 and Z_2 are jointly Gaussian zero mean r.v.s with variance N and covariance ρN where $0 < \rho < 1$. Your goal now is to optimally detect B from $\mathbf{Y} = [Y_1\ Y_2]^T$. Similarly, let $\mathbf{Z} = [Z_1\ Z_2]^T$ and let $\hat{B}(\mathbf{Y})$ be the estimate that minimizes $\mathbb{P}(B \neq \hat{B}(\mathbf{Y}))$.

b. (2 points) You decide to first whiten the noise in your observations by multiplying the observation vector $\mathbf{Y}$ by a matrix A such that $A\mathbf{Z}$ has identity covariance matrix. Let $\tilde{\mathbf{Y}} = A\mathbf{Y}$ and $\hat{B}(\tilde{\mathbf{Y}})$ be the optimal estimate of B from $\tilde{\mathbf{Y}}$, i.e. it minimizes $\mathbb{P}(B \neq \hat{B}(\tilde{\mathbf{Y}}))$. Are $\mathbb{P}(B \neq \hat{B}(\mathbf{Y}))$ and $\mathbb{P}(B \neq \hat{B}(\tilde{\mathbf{Y}}))$ equal to each other? Make a rigorous argument.

c. (2 points) Verify that the following matrix can be used to whiten the noise:

$$A = \frac{1}{\sqrt{N(1-\rho^2)}} \begin{bmatrix} \sqrt{1-\rho^2} & 0 \\ -\rho & 1 \end{bmatrix}.$$

d. (2 points) Write down the decoding rule $\hat{B}(\mathbf{Y})$. Your answer should be in terms of Y_1, Y_2, N, ρ and P. You can also depict the decision regions by a properly labeled picture if you find it easier. There is no need to reduce the decoding rule to its simplest form.

e. (2 points)What is the probability of error for $\hat{B}(\mathbf{Y})$? How does it compare to the probability of error in part (a)? Explain the improvement.

f. (3 points)Now you want to see if you can improve your communication system by using a different pair of codewords of the same power. Instead of transmitting $\mathbf{X} = [\sqrt{P}\,0]^T$ when $B = 1$ and $\mathbf{X} = [-\sqrt{P}\ 0]^T$ when $B = 0$ over the two consecutive uses of the channel, you decide to transmit $\mathbf{X} = [a\ b]^T$ when $B = 1$ and $\mathbf{X} = -[a\,b]^T$ when $B = 0$ where $a^2 + b^2 = P$. Find a and b that minimize the probability of error.

g. (2 points) What is the probability of decoding error when you use the a and b you found in the previous part and the corresponding optimal decoding rule at the receiver? How do you explain the improvement with respect to part (e)?

3. Exercise 8.7 from the text (10 points)

Note the following corrections on the book have been made below:

- In part (c), you need to find $f_{Y_1,Y_2|X,\phi}(y_1, y_2 \mid 0, 0)$, not $p_{Y_1,Y_2|X,\phi}(y_1, y_2 \mid 0, 0)$.
- In part (d), it should be $\Pr\{V_1 > y_1^2 + y_2^2 \mid X = 0\}$ instead of $\Pr\{V_1 > y_1^2 + y_2^2\}$.

Binary frequency shift keying (FSK) with incoherent reception can be modeled in terms of a four-dimensional observation vector $\mathbf{Y} = (Y_1, Y_2, Y_3, Y_4)^\top$, where $\mathbf{Y} = \mathbf{U} + \mathbf{Z}$. The L-dimensional rv $\mathbf{Z} \sim \mathcal{N}(0, \sigma^2 I)$ is independent of X. Under $X = 0$, $\mathbf{U} = (a\cos\phi, a\sin\phi, 0, 0)^\top$, whereas under $X = 1$, $\mathbf{U} = (0, 0, a\cos\phi, a\sin\phi)^\top$. The rv ϕ is uniformly distributed between 0 and 2π and is independent of X and $\mathbf{Z}$. The a priori probabilities are $p_0 = p_1 = 1/2$.

a) (2 points) Convince yourself from the circular symmetry of the situation that the ML receiver calculates the sample values v_0 and v_1 of $V_0 = Y_1^2 + Y_2^2$ and $V_1 = Y_3^2 + Y_4^2$ and chooses $\hat{x} = 0$ if $v_0 \geq v_1$ and chooses $\hat{x} = 1$ otherwise.

b) (2 points) Find $\Pr\{V_1 > v_1 \mid X = 0\}$ as a function of $v_1 > 0$.

c) (2 points) Show that

$$f_{Y_1,Y_2|X,\phi}(y_1, y_2 \mid 0, 0) = \frac{1}{2\pi\sigma^2} \exp\left[\frac{-y_1^2 - y_2^2 + 2y_1 a - a^2}{2\sigma^2}\right].$$

d) (2 points) Show that

$$\Pr\{V_1 > V_0 \mid X = 0, \phi = 0\} = \int f_{Y_1,Y_2|X,\phi}(y_1, y_2 \mid 0, 0) \Pr\{V_1 > y_1^2 + y_2^2 \mid X = 0\} dy_1 dy_2,$$

and show that this is equal to $(1/2)\exp\left(-a^2/(4\sigma^2)\right)$.

e) (2 points) Explain why this is the probability of error (i.e., why the event $V_1 > V_0$ is independent of ϕ), and why $\Pr\{e \mid X = 0\} = \Pr\{e \mid X = 1\}$.

4. BEC and BSC (13 points)

Consider the following communication problem. There are two equiprobable hypotheses. Under hypothesis $H = 0$, the transmitter sends $\bar{c}_0 = (0, 0, \ldots, 0)^T$ and under $H = 1$ it sends $\bar{c}_1 = (1, 1, \ldots, 1)^T$, both of length n, where n is some positive odd integer. The transmitted codeword $\bar{X} \in \{\bar{c}_0, \bar{c}_1\}$ is then passed through two independent discrete memoryless channels with the outputs denoted by $\bar{Y}_1$ and $\bar{Y}_2$ respectively. The receiver uses the outputs of both channels $\bar{Y}_1$ and $\bar{Y}_2$ to optimally decode the transmitted message.

The first channel is a Binary Erasure Channel (BEC) with erasure probability $p \in [0, 1]$, as depicted in Fig. 1: it has an input alphabet $\mathcal{X} = \{0, 1\}$, an output alphabet $\mathcal{Y}_1 =$

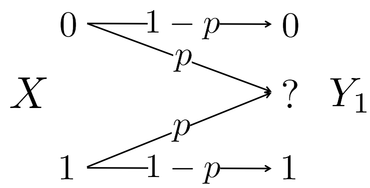

Figure 1: Binary erasure channel.

$\{0, 1, \text{“?”}\}$ and a transition probability $p_{Y_1|X}(y_1|x)$ given by

$$p_{Y_1|X}(y_1|x) = \begin{cases} 1 - p & \text{if } y_1 = x \\ p & \text{if } y_1 = \text{“?”} \end{cases}$$

The channel is discrete memoryless in the sense that when we use it to transmit an n-tuple $\bar{x}$, the ith output only depends on the ith input, i.e.,

$$p_{\bar{Y}_1|\bar{X}}(\bar{y}_1|\bar{x}) = \prod_{i=1}^{n} p_{Y_1|X}(y_{1i}|x_i).$$

The second channel is a Binary Symmetric Channel (BSC) with crossover probability $q \in [0, 1/2]$, as depicted in Fig. 2: it has an input alphabet $\mathcal{X} = \{0, 1\}$, an output

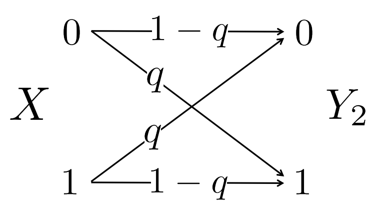

Figure 2: Binary symmetric channel.

alphabet $\mathcal{Y}_2 = \{0, 1\}$ and a transition probability $p_{Y_2|X}(y_2|x)$ given by

$$p_{Y_2|X}(y_2|x) = \begin{cases} 1 - q & \text{if } y_2 = x \\ q & \text{otherwise} \end{cases}$$

The channel is also discrete memoryless in the same sense as above.

(a) Assume $p = 0$ for the BEC.

  (i) (1 point) Describe the MAP rule for deciding whether $H = 0$ or $H = 1$ from $(\bar{Y}_1, \bar{Y}_2)$.

  (ii) (1 point) Determine the associated probability of error P_e.

(b) Assume $p = 1$ for the BEC.

  (i) (2 points) Identify a scalar integer valued sufficient statistic from $(\bar{Y}_1, \bar{Y}_2)$ for decoding the transmitted message.

  (ii) (1 point) Describe the MAP rule in terms of this sufficient statistic.

  (iii) (2 points) Determine the associated probability of error P_e. Hint: You can leave the result in terms of a summation.

(c) Consider the general case $p \in [0, 1]$.

  (i) (2 points) Identify a simple sufficient statistic from $(\bar{Y}_1, \bar{Y}_2)$ for decoding the transmitted message.

  (ii) (2 points) Describe the MAP rule in terms of this sufficient statistic.

  (iii) (3 points)Determine the associated probability of error P_e. Hint: You can leave the result in terms of a summation.