# Homework 1

Due on Wednesday, **3/26/2018**, at **10:00 AM** in class.

Be noted that late homework will **NOT** be accepted!!

## Grading Policy

- Please work all the four problems. However, we will randomly pick only three of them to grade.
- Each subproblem takes up to 10 points. The total of the to-be-graded three problems is 60 points.
- You need to provide detailed proof or derivations that lead to your final answers. Add comments to your programming codes. You will receive no points if your answers are not supported by detailed explanations. So, do not skip steps.
- Partial points will be credited to you when a wrong answer is accompanied by detailed correct reasoning.
- Discussions are encouraged. But plagiarism is strictly prohibited. **You will fail the course as penalty for plagiarism!!**

## Part I -- Reading assignment

Please read Section 1.1∼Section 1.5, Chapter 2 (except Sec. 2.3.8), and Chapter 3 of the textbook. You don't need to turn in anything for this part. I didn't touch Section 1.5 during class lectures, but the materials therein will be required in later chapters. For those of you who haven't taken Stochastic Processes or Detection & Estimation, please familiarize yourself to Section 1.5 by self-study.

## Part II -- Problem assignment

1. (a) Prove the following Bayes' rule:

$$f_{Y|X}(y|x) = \frac{P[X=x|Y=y]f_Y(y)}{\int_y P[X=x|Y=y]f_Y(y)dy},$$

   where $Y$ is a continuous random variable with pdf $f_Y(y)$ and $X$ is a discrete random variable. If we treat $X$ as the data and $Y$ the unknown parameter, then $f_{Y|X}(y|x)$ is the posterior density whereas $P[X=x|Y=y]$ can be thought of as the likelihood function.

   (b) Apply the Bayes' rule in Part (a) to the following problem:
   Suppose you are given a coin without knowing its bias (probability of heads). You first model the bias as a continuous random variable $Y$ uniformly distributed between 0 and 1. To get more information about the bias of the coin, you perform repeated trials by flipping the coin independently 10 times. Let $X$ be the number of heads you get. Suppose that the result of the repeated trials gives $X = 8$. Please find the conditional density $f_{Y|X}(y|8)$ of $Y$ given the resulted data $X = 8$. Also find the value of $y$ that maximizes the conditional density $f_{Y|X}(y|8)$.

2. Let $X_1$, $X_2$ and $X_3$ be jointly Gaussian random variables with the following properties: $X_1 \sim \mathcal{N}(0,1)$ and $X_2 \sim \mathcal{N}(0,1)$ are independent, $X_3 \sim \mathcal{N}(0,10)$, $E[X_1X_3] = -1$, and $E[X_2X_3] = 2$.

(a) Show that the covariance matrix of $[X_1,\ X_2]^T$ conditioning on $X_3$ takes the form

$$\mathbf{K} = \mathbf{I} + \alpha \cdot \mathbf{u}\mathbf{u}^T, \tag{1}$$

where $\mathbf{I}$ is the $2 \times 2$ identity matrix, $\mathbf{u} \in \mathbb{R}^{2\times 1}$ with $||\mathbf{u}|| = 1$, and $\alpha \in \mathbb{R}$ is a constant. What are the values of the vector $\mathbf{u}$ and the constant $\alpha$ in this problem?

(b) Show that in general any covariance matrix that takes the form in equation (1) must have $\mathbf{u}$ as an eigenvector.

3. (a) Consider bivariate (2-D) Gaussian random variables $X$ (with mean $\mu_X$ and variance $\sigma_X^2$) and $Y$ (with mean $\mu_Y$ and variance $\sigma_Y^2$). Show that the 2-D bivariate Gaussian pdf provided by equation (2.43) of the textbook can be re-written as

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \cdot e^{-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right\}},$$

where $\rho \triangleq \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X\sigma_Y}$ is the correlation coefficient of $X$ and $Y$.

(b) Consider the following pdf for $X$ and $Y$:

$$f_{XY}(x,y) = \frac{1}{\pi\sqrt{3/4}} \cdot e^{-\frac{1}{2}\left(\frac{4}{3}x^2 + \frac{16}{3}y^2 + \frac{8}{3}xy - 8x - 16y + 16\right)}. \tag{2}$$

1) Find $E[X], E[Y], \mathrm{Var}(X), \mathrm{Var}(Y)$, and $\mathrm{Cov}(X,Y)$, where $\mathrm{Cov}(X,Y) \triangleq E[(X-\mu_X)(Y-\mu_Y)]$.

2) Plot the pdf (in equation (2)) and the contour of $f_{XY}(x,y)$ at $f_{XY}(x,y) = 10^{-2}$ using MATLAB (or other software tools you prefer). (You can use the functions `mesh` and `contour` in MATLAB)

4. It is known that a certain type of heart disease is highly correlated with the patient's gender, type of chest pain, and cholesterol level. In this problem, you will implement the maximum a posteriori probability (MAP) detector of the heart disease for 50 patients, with their gender, type of chest pain, and cholesterol level provided in the test data set `heartdatasetTesting.mat` (which can be found as a separate file together with this homework assignment). The mathematical formulation is as follows:

Let $H, S, C$ be discrete random variables such that

$$H = \begin{cases} 0, & \text{if the patient does not have a heart disease,} \\ 1, & \text{if the patient has a heart disease,} \end{cases}$$

$$S = \begin{cases} 1, & \text{if the patient is a female,} \\ 2, & \text{if the patient is a male,} \end{cases}$$

$$C = \begin{cases} 1, & \text{if the pain is typical angina,} \\ 2, & \text{if the pain is atypical angina,} \\ 3, & \text{if the pain is non-anginal pain,} \\ 4, & \text{if without chest pain symptoms.} \end{cases}$$

Let $X$ be the continuous random variable denoting the patient's serum cholesterol in mg/dl.

In order to minimize the probability of detection error, you need to implement the MAP detection rule to determine whether a patient has the disease based on his/her gender, chest pain type, and cholesterol level (See Section 1.5 for details of MAP). More specifically, the MAP rule is

$$P[H=1|S=s,C=c,X=x] \underset{\hat{H}=0}{\overset{\hat{H}=1}{\gtrless}} P[H=0|S=s,C=c,X=x],$$

where $\hat{H}$ is the decision result, $s \in \{1,2\}$ and $c \in \{1,2,3,4\}$. The above MAP decision rule compares the posterior probabilities $P[H=h|S=s,C=c,X=x]$, $h \in \{0,1\}$: if $P[H=1|S=s,C=c,X=x] > P[H=0|S=s,C=c,X=x]$, then the detector produces the decision of $\hat{H}=1$; otherwise the decision goes to $\hat{H}=0$.

To evaluate the posterior probabilities, you need to learn relevant likelihood functions and prior distribution from the training data set `heartdatasetTraining.mat`. This part has been written in the MATLAB code `heartDisease.m`. You need to extend the code to solve the following two problems.

(a) Assume that $S, C$, and $X$ are conditionally independent given $H$. We further assume that $f_{X|H}(x|1) \sim \mathcal{N}(254, 2047)$, $f_{X|H}(x|0) \sim \mathcal{N}(245, 2182)$. From the given MATLAB code `heartDisease.m`, the values of $p_{S|H}[s|h]$ and $p_{C|H}[c|h]$ for all $s$, $c$, and $h \in \{0,1\}$ can be empirically evaluated using the provided data, which are required when finding the posterior probability $P[H=h|S=s,C=c,X=x]$.
Please use the information provided in the code `heartDisease.m` and extend the code to produce the MAP detection result $\hat{H}(s,c,x)$ for the 50 patients with their $S=s$, $C=c$, $X=x$ given in the test data `heartdatasetTesting.mat`. In other words, you need to write a MATLAB code based upon `heartDisease.m` to implement the MAP detection, and output the detection result of the 50 patients. You have to clearly describe how you obtain the posterior probability in your written homework sheet and add corresponding comments in your code.

(b) Following part (a), please evaluate the detection error rate of the MAP detector by comparing your result in part (a) to the true $H$ of each patient in the test data `heartdatasetTesting.mat`.