Homework 1

❗ For each problem, please clearly show your reasoning and write all the steps.
G As data scientists, you should feel free to google it whenever you see something unfamiliar.
☺ Group discussion for the homework is encouraged, but you have to write your answer by yourself. Also, you are always welcome to discuss the problems with me.

Task 0.

Read the relevant chapters in the textbooks listed on Sakai.

Problem 1. (5pt)

In this problem, we will look at a concrete example of Bayesian decision. This problem is adapted from Ex. 2.2 in [HaTF] (Ah, you found yourself a lazy instructor...).

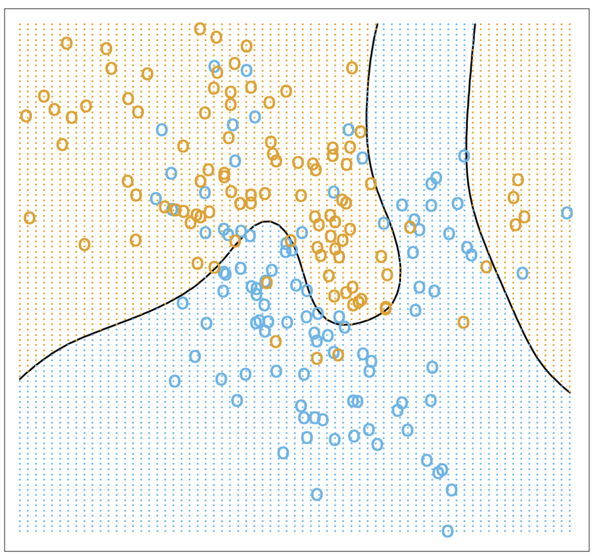

Figure 1: Simulation results of the experiment.

Consider the following experiment. Firstly a color in {BLUE, ORANGE} is randomly decided with $p(\text{BLUE}) = p(\text{ORANGE}) = 1/2$. If it is BLUE, then a point is drawn randomly from $\mathcal{N}([1, 0]^{\mathrm{T}}, \boldsymbol{I})$ where $\boldsymbol{I}$ is the identity matrix in $\mathbb{R}^2$; if it is ORANGE, then a point is drawn randomly from $\mathcal{N}([0, 1]^{\mathrm{T}}, \boldsymbol{I})$. Fig. 1 illustrates the simulation result of such an experiment. Suppose you know the above information. Write the equation of the Bayesian decision boundary.

Problem 2. (10pt)

In this problem, we will derive the MLE for geometric distributions.
A geometric distribution models the number of Bernoulli trials (e.g. coin tossing) needed to get one success. Specifically, let X be a random variable. $X \sim \text{Geom}(\theta)$ if for any ingeter $x \geq 1$,

$$P(X = x) = \theta(1-\theta)^{x-1} .$$

1. Find $\mathbb{E}X$ and $\text{Var}(X)$ if $X \sim \text{Geom}(\theta)$.
2. Consider the sample $\mathcal{X} = \{x_n\}_{n=1}^N$ where $x_n \sim^{i.i.d.} \text{Geom}(\theta)$. For the parameter θ above, write the likelihood $l(\theta|\mathcal{X})$ and the log-likelihood $\mathcal{L}(\theta|\mathcal{X})$.
3. Find the maximum likelihood estimator $\hat{\theta}_{\text{MLE}}$.
4. Is $\hat{\theta}_{\text{MLE}}$ biased?

Problem 3. (5pt+5pt(bonus))

In this problem, we will discuss the conjugate priors for Gaussian likelihoods. For simplicity, we consider the scalar case.

1. Suppose the prior
$$p(\theta) \sim \mathcal{N}(0, \sigma_0^2)$$
and the likelihood
$$p(x|\theta) \sim \mathcal{N}(\theta, \sigma^2)$$
where both σ_0 and σ are known. Show that $p(\theta)$ is a conjugate prior for $p(\mathcal{X}|\theta)$ where $\mathcal{X} = \{x_n\}_{n=1}^N$.
2. (bonus) The *precision* of a Gaussian model is defined to be $\beta = \sigma^{-2}$ (see, e.g. Ch. 16.3.2 of Alpaydin's textbook, or Ch. 1.2.5 of Biship's PRML textbook). Consider the likelihood
$$p(x|\beta) = \sqrt{\frac{\beta}{2\pi}} \exp\left(-\frac{\beta x^2}{2}\right) .$$
Show that the Gamma distribution
$$p(\beta) \sim \text{gamma}(a, b) = \frac{b^a}{\Gamma(a)} \beta^{a-1} \exp(-b\beta)$$
is a conjugate prior for $p(\mathcal{X}|\beta)$ where $\mathcal{X} = \{x_n\}_{n=1}^N$.

Problem 4. (10pt)

1. Show the following results.
   (a) $\text{tr}(\boldsymbol{AB}) = \text{tr}(\boldsymbol{BA})$ where tr denotes the trace of a matrix, $\boldsymbol{A}$ is an $m \times n$ matrix and $\boldsymbol{B}$ is an $n \times m$ matrix.
   (b) $\dfrac{\partial \text{tr}(\boldsymbol{AB})}{\partial \boldsymbol{A}} = \boldsymbol{B}^{\text{T}}$ where $\boldsymbol{A}$ is an $m \times n$ matrix and $\boldsymbol{B}$ is an $n \times m$ matrix.
   (c) $\dfrac{\partial \log(|\boldsymbol{A}|)}{\partial \boldsymbol{A}} = (\boldsymbol{A}^{-1})^{\text{T}}$ where $|\cdot|$ denotes the determinant of a matrix and $\boldsymbol{A}$ is an $m \times m$ matrix.
2. Derive the maximum likelihood estimates for the mean vector and the covariance matrix of a multivariate Gaussian density, given the sample $\mathcal{X} = \{\boldsymbol{x}_n\}_{n=1}^N$, where each $\boldsymbol{x}_n \in \mathbb{R}^d$. You may want to use the results in Part 1.

Problem 5. (10pt)

Write a journal about what you have learned during this week. You can think about e.g. the following questions: What have you learned this week? Which topic is the most interesting? Which topic is the most difficult? What application is useful? What techniques have you mastered during the week? What can the instructor do to improve the learning process? What have you learned from your classmates?