EECE5644 – 2018 Fall – Exam 1

Please submit your answers in a single file named as FirstLast_Exam1.pdf to the Assignments section of this course on Blackboard by 5pm ET on Wednesday, October 17, 2018.

In the pdf file name, FirstLast refers to your given and family names; this will help us easily identify the owner of the submitted pdf when downloaded from BB for grading (thanks).

If you have any questions, please send them to…

Brady section members: d.brady@northeastern.edu

Erdogmus section members: eece5644@googlegroups.com

Academic Honesty

By submitting your solutions for this take-home exam you attest that all of the work presented in your submission is the result of your own work.

Violations of Northeastern's academic honesty code of the university will result in actions that are consistent with university policy and procedures.

In fewer words, note that this is an exam and working together is absolutely not-allowed.

Problem 1 (15 points)

A random variable K~Poisson(λ) can take nonnegative integer values {0, 1, 2, …}, and the probability of observing the value *k* from this set is $P(K = k) = e^{-\lambda}\lambda^k/k!$ where λ>0 is the model parameter. Note that the mean and variance for this Poisson distribution are $E[K] = \lambda$ and $Var[K] = \lambda$.

We are given *N* independent identically distributed samples $D = \{k_1, k_2, \dots, k_N\}$ that are thought to be generated by this Poisson(λ) random process.

a) Parameter Estimation (5 points)

- Determine the maximum likelihood (ML) estimate of the parameter $\hat{\lambda}_{ML}$, as well as the mean $E[\hat{\lambda}_{ML}]$ and the variance $Var[\hat{\lambda}_{ML}]$ of this estimator;
- Determine the maximum a posteriori (MAP) estimate of the parameter $\hat{\lambda}_{MAP}$, assuming a Gamma(α, β) prior for the model parameter, as well as the mean $E[\hat{\lambda}_{MAP}]$ and the variance $Var[\hat{\lambda}_{MAP}]$ of this estimator. The prior for the parameter is specifically

$$p(\lambda; \alpha, \beta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda} & \text{for } \lambda \geq 0 \\ 0 & \text{for } \lambda < 0 \end{cases}.$$

b) Classification (10 points)

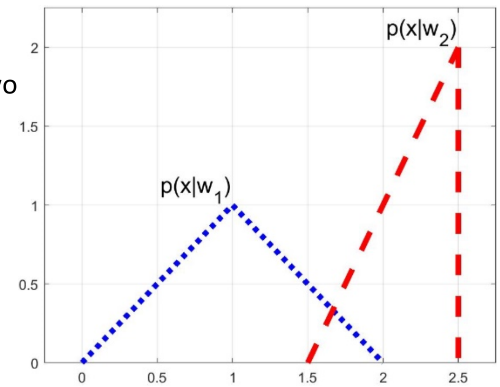
Consider *M* classes with prior probabilities $p_1, \dots, p_M$ where the feature probability distribution for class m is Poisson(λ_m) with distinct parameters for all $m \in \{1, \dots, M\}$. Determine the MAP (0-1 loss) decision rule to classify dataset D specified above, where all samples come from the same class. Simplify the discriminant expressions as much as possible.

Problem 2 (25 points)

Consider a feature x that has probability distributions under two competing class labels as given/shown.

$$p(x|w_1) = \begin{Bmatrix} x & if\ 0 \le x \le 1 \\ 2-x & if\ 1 \le x \le 2 \\ 0 & otherwise \end{Bmatrix}$$

$$p(x|w_2) = \begin{Bmatrix} 2x-3 & if\ 3/2 \le x \le 5/2 \\ 0 & otherwise \end{Bmatrix}$$



(a) (10 points) Given loss values $Loss(Deciding\ w_i\ given\ truth\ is\ w_j) = \lambda_{ij} \ge 0\ for\ i,j \in \{1,2\}$ and class prior probabilities $p(w_1)\ and\ p(w_2)$, express the minimum expected loss decision rule using a discriminant function that is simplified as much as possible. Please show all steps during your simplification process. Also indicate the particular minimum expected loss decision rule corresponding to the case where 0-1 loss is used, and class priors are equal (maximum likelihood classifier).

(b) (15 points) Determine analytically and sketch the ROC curve using $P(Decision = w_2|Truth = w_2)$ for the vertical axis and $P(Decision = w_1|Truth = w_1)$ for the horizontal axis. Indicate the point on the ROC curve that corresponds to the maximum likelihood classifier.

Problem 3 (30 points)

Implement code that accomplishes the tasks described below. Submit requested results and the code that generated those results.

Implement Code To Generate Samples: Generate N iid random 2-dimensional samples from two Gaussian pdfs $\mathcal{N}(\mu_i, \Lambda_i)$ with specified prior (class) probabilities $p(w_0)$ and $p(w_1)$, where

$$\mu_0 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \quad \mu_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \Lambda_0 = \begin{bmatrix} 16 & 0 \\ 0 & 1 \end{bmatrix}, \quad \Lambda_1 = \begin{bmatrix} 1 & 0 \\ 0 & 16 \end{bmatrix}.$$

Make sure to keep track of the true label of each sample.

Implement Code for the Minimum Expected Loss Classifier: Determine and implement the minimum expected loss classifier parametrized by a threshold γ in the following form:

$$ln p(x; \mu_1, \Lambda_1) - ln p(x; \mu_0, \Lambda_0) \underset{Decide\ 0}{\overset{Decide\ 1}{\gtrless}} ln\gamma$$

where γ>0 is a scalar that depends on loss and prior values. For different values of this threshold, classify each sample, empirically estimate the true positive and false positive probabilities (referring to class 1 as positive and class 0 as negative here).

Implement Code for the Fisher LDA Classifier: Implement the Fisher LDA classifier using the true means and covariances to obtain a classifier in the form

$$w^T x \underset{Decide\ 0}{\overset{Decide\ 1}{\gtrless}} \tau$$

and like the previous classifier, for different values of the threshold $\tau \in (-\infty, \infty)$ empirically estimate the true and false positive probabilities.

<u>Report the following:</u> Generate N=10000 iid samples from the two Gaussians specified above with priors $p(w_0)$ = 0.35 and $p(w_1)$ = 0.65. Using the two classifier codes, process this data set to make classification decisions for various appropriate sets of threshold values and in a single plot, present in an overlaid fashion the ROC curves for the two classifiers empirically estimated using this data set. (2x10 points)

On the ROC curve of the minimum expected loss classifier, specifically indicate the point that corresponds to 0-1 loss (MAP classifier). For the MAP classifier, report the probability of error estimated empirically by counting the erroneously classified samples (divided by the total number of samples). (5 points)

Similarly, on the ROC curve of the Fisher LDA classifier, specifically indicate the point that corresponds to a threshold that achieves minimum probability of error, estimated empirically in the same fashion as in the other case. Report this smallest empirical probability error value achieved by LDA. (5 points)

Please include some brief explanation for all plots and numerical values reported. Also include all your code in the pdf along with these results.

Problem 4 (30 points)

A vehicle is to be localized in 2-dimensional space based on distance measurements from *K* reference positions. These distance measurements are $r_i = d_i + n_i$ for $i = 1, \dots, K$, where d_i is the true distance between the vehicle and the *i*th reference point, and n_i is a zero mean Gaussian distributed measurement noise with known variance σ_i^2. The noise in each measurement is independent from the others.

Assume that we have the following prior knowledge regarding the position of the vehicle:

$$p\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{1}{2}[x \quad y]\begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix}^{-1}\begin{bmatrix} x \\ y \end{bmatrix}}$$

In your answer, please use $\begin{bmatrix} x_T \\ y_T \end{bmatrix}$ to indicate the true position of the vehicle, $\begin{bmatrix} x_i \\ y_i \end{bmatrix}$ to indicate the position of the *i*th reference point, and $\begin{bmatrix} x \\ y \end{bmatrix}$ to indicate a candidate vehicle position under consideration.

(10 points) Express the optimization problem that needs to be solved to determine the MAP estimate of the vehicle position. Simplify the objective function so that exponentials and additive/multiplicative terms that do not depend on the candidate vehicle position are eliminated as much as possible (hint: do not expand squared-parentheses; you may stop at that stage so that your answer looks like a sum of squared error terms plus a quadratic term from the prior).

(20 points) Please implement the following as code.

Set the true vehicle location to be inside the circle with unit radius centered at the origin. For each value of *K* between 1 and 4 repeat the following.

Place evenly spaced *K* landmarks on a circle with unit radius centered at the origin. Set measurement noise standard deviation to 0.3 for all range measurements. Generate *K* range measurements according to the model specified above (if a range measurement turns out to be negative reject it and resample; all range measurements need to be nonnegative).

Plot the equilevel contours of the MAP estimation objective for the range of horizontal and vertical coordinates from -2 to +2; superimpose the true location of the vehicle on these equilevel contours (e.g. use a + mark), as well as the landmark locations (e.g. use a o mark for each one).

Provide plots of the MAP objective function contours for each value of K. When preparing your final contour plots for different *K* values make sure to plot contours at the same function value across each of the different contour plots for easy visual comparison of the MAP objective landscapes (e.g. in the sample code we prepared as a solution contours at levels 10^z where z takes 5 values linearly spaced between 3/2 and 5/2 to provide 5 logarithmically spaced cross sections of the MAP objective worked well).

Supplement your plots with a brief description of how your code works. Comment on the behavior of the MAP estimate of position (visually assessed from the contour plots; roughly the center of the innermost contour) relative to the true position. Does the MAP estimate get closer to the true position as K increases? Does it get more certain? Explain how your contours justify your conclusions.

Please include some brief explanation for all plots and numerical values reported. Also include all your code in the pdf along with these results.