EECE5644 Fall 2019 – Homework 1

Start: Monday, 2019-September-23
Submit: Monday, 2019-September-30 before 09:00ET

Please submit your solutions to Blackboard in a single PDF file that includes all math and numerical results in the main body (please start a new page for each question). Make sure that you cite all resources you benefit from (books, papers, software packages you use) as appropriate in your submissions. Also include your code that is used to generate numerical results for your answers in one of the following ways: (Undesirable) copy/paste into an appendix section of the PDF document, (Acceptable) upload a ZIP file containing all code files, (Preferred) keep your code in an online version control repository (such as GitHub, BitBucket, etc) and as part of your answer to each question that requires code, provide a link to the relevant online repository.

Note that this is a graded assignment and the entirety of your submission must contain only your own work. You may use written or other forms of literature and software packages available to you, as long as these sources are properly acknowledged in your submission as cited references.

Question 1 (10%)

1. Let $x \in \mathbb{R}$ be a random variable with mean $E[x] = \mu$. Show that the variance of x, $E[(x-\mu)^2]$ is $E[x^2] - \mu^2$.
2. Let $\mathbf{x} \in \mathbb{R}^n$ be an n-dimensional random vector with mean $E[\mathbf{x}] = \mu$. Show that the covariance matrix of $\mathbf{x}$, $E[(\mathbf{x}-\mu)(\mathbf{x}-\mu)^{\mathrm{T}}]$ is $E[\mathbf{x}\mathbf{x}^{\mathrm{T}}] - \mu\mu^{\mathrm{T}}$.

Question 2 (20%)

Suppose that a real-valued measurement x is generated by one of two equally probable probability density functions (pdfs), both of the parametric form $p(x|L=l) \propto e^{-|x-a_l|/b_l}$ for $l \in 1,2$ and $b_l > 0$ for both pdfs.

1. Find the normalization factor for this parametric pdf family as a function of the respective a and b parameters, such that these pdfs integrate to 1 along the real axis.
2. Determine the simplest analytical expression for the log-likelihood ratio between class labels 1 and 2 evaluated at a given x, $\ell(x) = lnp(x|L=1) - lnp(x|L=2)$.
3. Generate a plot of this log-likelihood-ratio function for the case $a_1 = 0, b_1 = 1$ and $a_2 = 1, b_2 = 2$ using a suitable programming language. Label the axes of the plot properly as $\ell(x)$ and x using legible font sizes, and include an intelligible title and caption for the plot. In all future assignments, this will be expected without explicit instruction. All figures must be properly labeled, titled, and captioned. If there are multiple curves or data visualizations in the plot also include a legend to help identify each component.

Question 3 (20%)

Consider a two-class setting where class prior probabilities are $P(L=1) = P(L=2) = 1/2$. We measure a one dimensional real-valued feature variable x, which takes values from $Uniform[a,b]$ for class 1, and $Uniform[r,t]$ for class 2. Here $a < r < b < t$, so the support of these two uniform class-conditioned feature pdfs p(x—L=1) and p(x—L=2) overlap. Derive a minimum probability of error classification rule.

Question 4 (30%)

Consider a two-class setting with equal class priors and a real-valued feature with Gaussian class-conditioned pdfs $\mathcal{N}(0,1)$ and $\mathcal{N}(\mu,\sigma^2)$.

1. Derive an expression for the classification/decision rule that achieves minimum probability of error. This will be a function of x that depends on parameters μ and σ. Please simplify the expression to a *nice* form.
2. For the case $\mu = 1$ and $\sigma^2 = 2$ generate plots that visualize the class-conditional pdfs $p(x|L = l)$ for $l \in 1,2$, as well as class posterior probabilities $p(L = l|x)$ for $l \in 1,2$. Demonstrate the decision boundary you found earlier in these visualizations.
3. Numerically estimate the achieved minimum probability of error for the specific case you visualized above. Describe your procedure and the numerical result intelligibly.
4. Consider the case where $\mu = 0$ and $\sigma \gg 1$ for the second class conditional pdf. Discuss what happens to the decision boundary and regions. Imagine and briefly describe a practical situation when this kind of data distribution might arise.

Question 5 (20%)

The pdf of an n-dimensional random vector drawn from $\mathcal{N}(\boldsymbol{\mu},\boldsymbol{\Sigma})$ is

$$p(\boldsymbol{\zeta}) = (2\pi)^{-n/2}|\boldsymbol{\Sigma}|^{-1/2}e^{-\frac{1}{2}(\boldsymbol{\zeta}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\zeta}-\boldsymbol{\mu})} \tag{1}$$

1. **Determine the pdf of x:** Let $\mathbf{x} = \mathbf{A}\mathbf{z} + \mathbf{b}$ be an n-dimensional random vector where $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathbf{b} \in \mathbb{R}^n$ are fixed and known, $\mathbf{z} \sim \mathcal{N}(\mathbf{0},\mathbf{I})$. Assuming that $\mathbf{A}$ is full rank, and by using/citing an appropriate theorem from probability theory, show that $\mathbf{x} \sim \mathcal{N}(\mathbf{b},\mathbf{A}\mathbf{A}^{\mathrm{T}})$.
2. **Determine Linear Transformation:** We want to generate a random vector drawn from $\mathcal{N}(\boldsymbol{\mu},\boldsymbol{\Sigma})$ using the linear transformation technique introduced and studies above. Determine $\mathbf{A}$ and $\mathbf{b}$ to achieve the desired mean vector and covariance matrix. Specifically, determine $\mathbf{A}$ and $\mathbf{b}$ in terms of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$.
3. **Implement Code:** Write code that takes in N, n, $\boldsymbol{\mu}$, and $\boldsymbol{\Sigma}$ and produces N samples of independent and identically distributed (iid) n-dimensional random vectors $\{\mathbf{x}_1,\ldots,\mathbf{x}_N\}$ drawn from $\mathcal{N}(\boldsymbol{\mu},\boldsymbol{\Sigma})$ using the linear transformation technique applied to samples of $\mathbf{z} \sim \mathcal{N}(\mathbf{0},\mathbf{I})$.