ECE C143A/C243A, Spring 2018 — **Homework #4**

Department of Electrical and Computer Engineering — Prof. J.C. Kao

University of California, Los Angeles — TAs T. Monsoor, X. Jiang and X. Yang

Due Wednesday, 16 May 2018, uploaded to Gradescope.
Covers material up to Discrete Classification II.
100 points total.

Please **box your final answers for each question.**

A probabilistic generative model for classification comprises class-conditional densities $P(\mathbf{y} \mid \mathcal{C}_k)$ and class priors $P(\mathcal{C}_k)$, where $\mathbf{y} \in \mathbb{R}^D$ and $k = 1, \ldots, K$. We will consider three different generative models in this problem set:

i) Gaussian, shared covariance

$$\mathbf{y} \mid \mathcal{C}_k \sim \mathcal{N}\left(\boldsymbol{\mu}_k,\, \Sigma\right)$$

ii) Gaussian, class-specific covariance

$$\mathbf{y} \mid \mathcal{C}_k \sim \mathcal{N}\left(\boldsymbol{\mu}_k,\, \Sigma_k\right)$$

iii) Poisson

$$y_i \mid \mathcal{C}_k \sim \text{Poisson}(\lambda_{ki})$$

In iii), y_i is the ith element of the vector $\mathbf{y}$, where $i = 1, \ldots, D$. This is called a *naive Bayes* model, since the y_i are independent conditioned on $\mathcal{C}_k$.

1. (20 points) Maximum likelihood (ML) parameter estimation
   In class, we derived the ML parameters for model i):

$$P(\mathcal{C}_k) = \frac{N_k}{N} \qquad \boldsymbol{\mu}_k = \frac{1}{N_k}\sum_{n\in\mathcal{C}_k} \mathbf{x}_n \qquad \Sigma = \sum_{k=1}^{K} \frac{N_k}{N} \cdot S_k,$$

   where

$$S_k = \frac{1}{N_k}\sum_{n\in\mathcal{C}_k} (\mathbf{x}_n - \boldsymbol{\mu}_k)(\mathbf{x}_n - \boldsymbol{\mu}_k)^{\mathrm{T}},$$

   N_k is the number of data points in class $\mathcal{C}_k$, and N is the total number of data points in the data set.

   (a) (10 points) Find the ML parameters for model ii), i.e., find: $P(\mathcal{C}_k)$, $\boldsymbol{\mu}_k$, Σ_k. (**Hint**: You should incorporate a Lagrange multiplier that $\sum_k P(\mathcal{C}_k) = 1$.)

   (b) (10 points) Find the ML parameters for model iii), i.e., find: $P(\mathcal{C}_k)$, λ_{ki}

2. (20 points) Decision boundaries
   In class, we derived the decision boundary between class $\mathcal{C}_k$ and class $\mathcal{C}_j$ for model i):

$$(\mathbf{w}_k - \mathbf{w}_j)^{\mathrm{T}}\mathbf{x} + (w_{k0} - w_{j0}) = 0,$$

   where

$$\mathbf{w}_k = \Sigma^{-1}\boldsymbol{\mu}_k \qquad w_{k0} = -\frac{1}{2}\boldsymbol{\mu}_k^{\mathrm{T}}\Sigma^{-1}\boldsymbol{\mu}_k + \log P(\mathcal{C}_k).$$

   For each of the models ii) and iii), we'll want to derive the decision boundary between class $\mathcal{C}_k$ and class $\mathcal{C}_j$ and say whether it is linear in $\mathbf{y}$.

   (a) (10 points) What is the decision boundary for model (ii)? Is it linear?

   (b) (10 points) What is the decision boundary for model (iii)? Is it linear?

3. (30 points) Simulated data
   please complete `hw4p3.ipynb`, which could be downloaded from CCLE, the data is `ps4_simdata.mat`
   Please print the Jupyter Notebook and attach to your HW.

4. (30 points) Real neural data
   please complete `hw4p4.ipynb`, which could be downloaded from CCLE, the data is `ps4_realdata.mat`
   Please print the Jupyter Notebook and attach to your HW.

   The neural data have been generously provided by the laboratory of Prof. Krishna Shenoy at Stanford University. The data are to be used exclusively for educational purposes in this course. Please print the Jupyter Notebook and attach to your HW.