CSC 480/580 Homework 4

Due: 4/30 (Tue)

Instructions:

- If you use math symbols, please define it clearly before you use it (unless they are standard from the lecture).
- You must provide the derivation for obtaining the answer and full source code for whatever problem you use programming. Please email your source codes to `csc480580@gmail.com`.
- Please use the problem & subproblem numbering of this document; do not recreate or renumber them.
- Submit your homework on time to gradescope. NO LATE DAYS, NO LATE SUBMISSIONS ACCEPTED.
- The submission must be one single PDF file (use Acrobat Pro from the UA software library if you need to merge multiple PDFs).
- Please include your answers to all questions in your submission to Gradescope. (Do not store your answers in your source codes or Jupyter notebooks - I will not look at them by default.)
  - You can use word processing software like Microsoft Word or LaTeX.
  - You can also hand-write your answers and then scan it. If you use your phone camera, I recommend using TurboScan (smartphone app) or similar ones to avoid looking slanted or showing the background.
  - Watch the video and follow the instruction: `https://youtu.be/KMPoby5g_nE` .
- Collaboration policy: do not discuss answers with your classmates. You can discuss HW for the clarification or any math/programming issues at a high-level.

Problem 1: Gaussian mixture models

Consider a Gaussian mixture model (GMM) on $K = 4$ components with the distribution over the i^{th} observation is given by,

$$z_i \sim \text{Categorical}(\pi)$$
$$x_i \mid z_i = k \sim \mathcal{N}(\mu_k, \Sigma_k).$$

Here z_i is a discrete assignment that controls which component the data x_i are generated from. The underlying model parameters are:

- $\pi_1 = 0.2$, $\mu_1 = (0, 0)$, $\Sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
- $\pi_2 = 0.3$, $\mu_2 = (4, 4)$, $\Sigma_2 = \begin{bmatrix} 1.5 & 1 \\ 1 & 1.5 \end{bmatrix}$
- $\pi_3 = 0.3$, $\mu_3 = (-4, -4)$, $\Sigma_3 = \begin{bmatrix} 2.5 & -2 \\ -2 & 2.5 \end{bmatrix}$
- $\pi_4 = 0.2$, $\mu_4 = (-8, 16)$, $\Sigma_4 = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}$

show all of your work for the following questions:

1. *(3pts) Plot the probability density function (PDF) contours of all four ground truth Gaussian components $\mathcal{N}(\mu_k, \Sigma_k), k = 1, \ldots, 4$ in one figure. For clarity, I recommend ignoring plotting contour lines for small PDF values around 0. You can do this by a Python command like (if using matplotlib)*

   ```
   plt.contour(X, Y, pdf, levels=np.linspace(np.min(pdf), np.max(pdf), 5)[1:])
   ```

2. *(3pts) Draw $n = 1000$ samples $S = (x_1, \ldots, x_{1000})$ from the GMM, and make a scatter plot based on the sample. Does it look like the contour plots in the previous question? Why?*
3. *(8pts) Implement the EM algorithm taught in the class to estimate the model parameters, and run it on S. Report the learned model parameters, and show the PDF contour plots of the learned Gaussian components, after 1, 3, 9 iterations, respectively.*