EE5601: Representation Learning, Fall 2019 (34)
Indian Institute of Technology Hyderabad
HW 1, Assigned: Tuesday 24.09.2019. 35 points
Due: Monday 30.09.2019 at 11:59 pm.

Note: *Please do not use built-in functions for the expectation maximization algorithm..*

1. Derive the expression for the optimal decorrelating linear transform for a set of observations $X \in R^{d \times N}$ where each row is assumed to be zero-mean. (5)

2. Derive the expressions for the partial derivates of the log likelihood function of a Gaussian Mixture Model (GMM) with respect to each of its parameters. Set these derivatives to zero and find the expressions for the "locally optimal" parameters in terms of the posterior probabilities and the observations. (10)

3. Implement the expectation maximization (EM) algorithm for estimating the parameters of a Gaussian Mixture Model (GMM). The GMM density is given by
$$p(\mathbf{x}|\boldsymbol{\theta}) = \sum_{k=1}^{K} \pi_k \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}_k, \Sigma_k),$$
where $\mathbf{x} \in R^d$, $\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \Sigma)$ is the multivariate Gaussian distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ. The parameter set $\boldsymbol{\theta} = [\pi_1, \ldots, \pi_K, \boldsymbol{\mu}_1, \ldots \boldsymbol{\mu}_K, \Sigma_1, \ldots, \Sigma_K]$. Your program must accept as inputs the observation matrix X (of size $d \times N$), and the mixture size K as inputs, and output the estimated parameter set $\hat{\boldsymbol{\theta}}$. Generate X on your own and experiment by varying your choices of $\boldsymbol{\theta}$. (20)