Problem 1: Assume that the scalar random variable $x$ is distributed according to the Gaussian mixture model

$$f(x) = 0.5N(5, 1) + 0.5N(-5, 1),$$

where $N(\mu, \sigma^2)$ denotes a Gaussian with mean $\mu$ and variance $\sigma^2$. a) Generate 500 realizations of $x$ that follow this pdf. Explain how you will generate the data. The samples $\{x_1, \ldots, x_{500}\}$ constitute your data.

Suppose now that you do not know exactly the previous pdf model but only that it follows the form

$$f(x) = wN(\mu_1, \sigma_1^2) + (1-w)N(\mu_2, \sigma_2^2),$$

where $\mu_1, \sigma_1^2, \mu_2, \sigma_2^2$ and $w \in (0, 1)$ parameters to be estimated. b) Recall the Expectation/Maximization (EM) procedure we presented in the class that estimates these parameters by introducing the latent probabilities $p_{j1}, p_{j2} = (1 - p_{j1})$. Probability $p_{j1}$ means that sample $j$ comes with probability $p_{j1}$ from the first Gaussian and with probability $p_{j2}$ from the second. Apply the EM to the data you have generated to estimate the required parameters. c) Consider now that you are given the exact values of $\sigma_1^2 = \sigma_2^2 = 1$ and it is not required to estimate them. Write the new form of the EM algorithm for this case and apply it to your data to estimate $\mu_1, \mu_2, w$. d) Repeat c) but this time assume that you are given the means $\mu_1 = 5, \mu_2 = -5$. Write the EM algorithm for $\sigma_1^2, \sigma_2^2, w$ and apply it to your data to estimate these parameters. e) Finally consider that $\mu_1, \mu_2$ are known but you are given the wrong values $\mu_1 = 5, \mu_2 = 0$. Estimate again $\sigma_1^2, \sigma_2^2, w$ using the EM algorithm from d). What do you observe when comparing d) to e)? What about when you compare b) to c)? f) In all estimation problems plot the log-likelihood ratio of your data to verify that it decreases with each iteration of the EM algorithm.

Problem 2: Let $X$ be a random vector of length 2 which is Gaussian with mean and covariance matrix given by

$$\mu = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \Sigma = 2 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + 0.1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. $$

a) Show that the previous model can be generated from the following equation by introducing the latent variables $z, W$

$$X = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} z + \sigma W$$

where $z$ is scalar Gaussian with mean 0 and variance 1, $W$ is a Gaussian vector of length 2, independent from $z$, with mean 0 and covariance matrix equal to the identity. Specify the exact values of $\mu_1, \mu_2, a_1, a_2, \sigma$ that will yield the original model. b) Using a) propose a method to generate realizations of $X$ and generate 500 such independent vectors. c) Assume now that the mean and the covariance matrix of your data is unknown and you would like to estimate them. c1) Use the classical method to make this estimate. c2) Observe that the covariance matrix has the special form we discussed in class and use SVD to make a more powerful estimate that takes into account the extra information about the special form of the covariance matrix. c3) Compare the estimates for the covariance matrix of the two methods by generating several 500-tuples and approximating the estimation error power of both methods.

We will discuss the problems tomorrow Monday, Nov. 26, at 6PM, in CBIM-22.
Your reports, in hard copy, must be submitted to Mr. Neelesh Kumar on Wednesday, Nov. 28, between 5:00-6:00PM, Room: CBIM-05.