## Homework # 3

1. Attached is a file containing the heights of men and women at Hope College, see file for details. Consider the two component Gaussian mixture model,

$$X = \begin{cases} \mathcal{N}(\mu_1, \sigma_1^2) & \text{with probability } p_1 \\ \mathcal{N}(\mu_2, \sigma_2^2) & \text{with probability } 1 - p_1 \end{cases} \tag{1}$$

where $\mathcal{N}(\mu, \sigma^2)$ is the normal distribution and $X$ models the height of a person when gender is unknown. Let $\theta = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, p_1)$.

   (a) The data file specifies gender, but pretend you don't have this information. Write down the log-likelihood function $\ell(\theta)$ and $\nabla\ell(\theta)$ given the height samples, i.e. in terms of $\hat{X}_i$. Write R (or Python) functions that calculate $\ell(\theta)$ and $\nabla\ell(\theta)$

   (b) Find the MLE for $\theta$ by

      i. Applying a steepest ascent iteration $\theta^{(i+1)} = \theta^{(i)} + s\nabla\ell(\theta)$.

      ii. Using **nlm** or an equivalent in Python.

   (c) Given your MLE in (b), use the distribution of $X$ to predict whether a given sample is taken from a man or woman. Intuitively, the two normal distributions of $X$ correspond to the male and female height distributions. Given a sample, $\hat{X}$, decide which normal the sample is most likely to come from and assign the gender accordingly. Determine what percentage of individuals are classified correctly.

2. (a) Let $Q$ be an $n \times n$ orthonormal matrix. Show that $Q^{-1} = Q^T$

   (b) Let

$$R = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \tag{2}$$

$$F = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3}$$

   (c) Show that $R$ rotates vectors by an angle $\theta$ and that $F$ reflects vectors about the $x$-axis

(d) Show that you can form any $2 \times 2$ orthonormal matrix using $R$ and $RF$.

3. Let $X \sim (\mu, \Sigma)$ where $\mu \in \mathbb{R}^n$ and $\Sigma$ is an $n \times n$ covariance matrix.

   (a) Let $M$ be an $n \times n$ invertible matrix. Show that $MX \sim (M\mu, M\Sigma M^T)$. (Hint: Write an integral expression for $P(MX \leq c)$ where the inequality means that the $i$th coordinate of $MX$ is less than the $i$th coordinate of $c$, which is a vector, for $i = 1, 2, \ldots, n$. Then, transform the integral using the change of variable $z = Mx$. Show that the resultant integral equals $P(Z \leq c)$ for $Z \sim N(\mu, M\Sigma M^T)$.)

   (b) Show that if $\Sigma$ is a diagonal matrix then the coordinates of $X$ are independent normals. (We did this in class. Here I want you to go through it yourself.

   (c) Write a function **MultiNorm($\mu$, $\Sigma$)** that samples from a multivariate normal with mean $\mu$ and covariance $\Sigma$. Your function can use **rnorm**, the univariate normal sampler in R, and the function **eigen** (or their equivalent in Python). (Hint: Recall the spectral decompostion $\Sigma = QDQ^T$. Consider $Q^T X$ and use 2a, 3a, 3b to construct the sampler). Don't just write the code, also explain why your sampler is correct.