Homework 2 – Due March 9

1. Let $\boldsymbol{X} \sim N_p(\boldsymbol{0}, \boldsymbol{I})$. Let Y be a random variable having the χ^2_p-density that is independent of $\boldsymbol{X}$. Let $\boldsymbol{Z} = \sqrt{Y}\frac{\boldsymbol{X}}{\|\boldsymbol{X}\|}$. Show that that the density of $\boldsymbol{Z}$ is also standard multivariate normal. [*10 points*]

2. Write an example in C to illustrate that a function passes its arguments *by value*. To do this, write a function which takes in two arguments: an integer and a pointer to an integer and then increments each by 1. Follow the location of the arguments inside and outside the function to illustrate the point. [*10 points*]

3. Write a function in C which illustrates the use of the matrix multiplication algorithm using CCS representation that you wrote in the previous assignment. [*20 points*]

4. Let $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n$ be a random multivariate sample such that $\boldsymbol{X}_i$ has only the first $1 \leq p_i \leq p$ coordinates that are observed. Assume that each $\boldsymbol{X}_i$ is a realization from the p_i-dimensional marginal distribution of $\mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Further, there are at least two is for which $p_i = p$. Answer the following questions:

   (a) Find the maximum likelihood estimator of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ using direct maximization of the loglikelihood. [*15 points*]

   (b) Use the expectation-maximization algorithm to formulate the maximum likelihood estimator of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. [*20 points*]

   (c) Compare the number of computations (floating point operations) needed in one EM step to the number of computations in the direct calculations. You may make simplifying assumptions as needed for calculating the number of operations. [*20 points*]

   (d) How do the results in (c) change if the p_i observed coordinates are not the first ones? [*5 points*]