# Homework 5

Due: Thursday, November 2, 2023, 1:00 pm on Gradescope

Please upload your answers timely to Gradescope. Start a new page for every problem. For the programming/simulation questions you can use any reasonable programming language. Comment your source code and include the code and a brief overall explanation with your answers.

1. (**12 pts**) Exercises 3.1 in text.

   **Exercise 3.1** **(a)** Let $X$, $Y$ be IID rv s, each with density $\mathsf{f}_X(x) = \alpha \exp(-x^2/2)$. In (b), we show that $\alpha$ must be $1/\sqrt{2\pi}$ in order for $\mathsf{f}_X(x)$ to integrate to 1, but in this part, we leave $\alpha$ undetermined. Let $S = X^2 + Y^2$. Find the probability density of $S$ in terms of $\alpha$.

   **(b)** Prove from (a) that $\alpha$ must be $1/\sqrt{2\pi}$ in order for $S$, and thus $X$ and $Y$, to be rv s. Show that $\mathsf{E}[X] = 0$ and that $\mathsf{E}\left[X^2\right] = 1$.

   **(c)** Find the probability density of $R = \sqrt{S}$. $R$ is called a *Rayleigh* rv.

2. (**10 pts**) Exercise 3.3 in text. Let $X$ and $Z$ be independent standard Gaussian random variables. Let $Y = |Z|\text{Sign}(X)$ where $\text{Sign}(X)$ is 1 if $X \geq 0$ and $-1$ otherwise.

   a) (**5 pts**) Show that $Y$ is also standard Gaussian.

   b) (**5 pts**) Is $Y + Z$ Gaussian? (Either prove that it is Gaussian or argue that it is not Gaussian.)

3. (**12 pts**) The moment generating function of a random vector $\mathbf{Z} \in \mathbb{R}^n$ is defined as

$$g_{\mathbf{Z}}(\mathbf{r}) = \mathbb{E}\left[\exp\left(\mathbf{r}^T \mathbf{Z}\right)\right]$$

   where $\mathbf{r} \in \mathbb{R}^n$ is an $n$-dimensional real vector. (The MGF may not exist for all $\mathbf{r} \in \mathbb{R}^n$ just as we discussed in the scalar case but we will soon see that for Gaussian random vectors, it exists everywhere.) Just like in the case of scalar random variables, the MGF fully characterizes the distribution of a random vector.

   a) (**3 pts**)Compute $g_{\mathbf{W}}(\mathbf{r})$ for $\mathbf{W} \sim \mathcal{N}(0, I_n)$, i.e. $\mathbf{W}$ is a standard Gaussian random vector.

   b) (**6 pts**)Assume $\mathbf{X}$ is a Gaussian random vector with covariance $K_{\mathbf{X}}$ and mean $\mu$. Show that

$$g_{\mathbf{X}}(\mathbf{r}) = \exp\left(\mathbf{r}^T \mu + \frac{\mathbf{r}^T K_{\mathbf{X}} \mathbf{r}}{2}\right)$$

c) (**3 pts**)Use the previous part to argue that if $\mathbf{X}$ and $\mathbf{Z}$ are Gaussian random vectors such that $K_{\mathbf{X}} = K_{\mathbf{Z}}$ and $\mu_{\mathbf{Z}} = \mu_{\mathbf{X}}$, then $\mathbf{X}$ and $\mathbf{Z}$ have the same distribution.

4. (**16 pts**) Download the data package at http://web.stanford.edu/class/ee278/homeworks/hw2-data.zip, where you'll find the MNIST handwritten digit dataset. The data package includes two image subsets (concerning digit 0 and digit 2) each containing 4999 instances of $28 \times 28$ pixels.

   (a) (**4 pts**) Each image can be vectorized as a 784-dimensional vector. Suppose that each image in subset $l$ are independently drawn from some distribution $\mathbb{P}_l$ $(l = 1, 2)$. Compute an estimate of the covariance matrix for each subset of images (i.e. the set of digit-0 images and the set of digit-2 images).

   (b) (**4 pts**) For each subset of images, compute and plot the first few eigenvectors (as images) of the estimated covariance matrix. These are the principal components of the data.

   (c) (**4 pts**) Compute an estimate of the covariance matrix for the entire set of images (including all digit-0 and digit-2 images). Plot its first few eigenvectors as images. Compare and contrast these plots with the ones you obtain in Part (b).

   (d) (**4 pts**) Plot the eigenvalues for the covariance matrices computed in Part (a) and Part (c). Explain your findings in the plots. For instance, from the eigenvalues, what can you say about the "effective" dimensionality of the data? What does your observation suggest for feature selection for say recognizing digits?

   Please include your code. You can use any programming language (e.g. Matlab, Python) for this problem, but you are not allowed to use off-the-shelf functions like "PCA" or "cov". For Matlab, you might find the functions *imshow* and *imread* useful in plotting and reading image data.