Geometric Foundations of Data Sciences
CS378(51240) CSE392(64409) M392C(54069) PHY341(55661)
Fall 2020: HW Exercise 1

Notes:

- Posted Aug. 31, 2020. Due before 9:00pm, Sept 14, 2020
- Students are allowed to discuss the homework problem sets with one another. However your **solutions must be written by yourself** and in your own words. Copying homework solutions from another student, from a textbook, from a webpage, or from any other source is not allowed.
- Please state the persons with whom the exercise solutions was discussed and any material that was consulted to help you solve the exercise.
- Please typeset your solutions if possible. If you submit handwritten solutions they must be written legibly. **The solution must be turned in via the Canvas assignment page**.

Exercise 1 [10pts]

Consider the matrix A and its singular value decomposition (SVD)

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3/5 & -4/5 & 0 \\ 0 & 4/5 & 3/5 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 4+\beta^2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4-\beta^2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{pmatrix},$$

where $\beta > 0$. Suppose $\|A\|_2 = 6$.

(a) [10pts] Solve for β, given that $\|A\|_2 = 6$.

(b) [10pts] Find the best rank-2 approximation matrix of A. Give out accurate final results with only one matrix.

Exercise 2 [10pts]

(1) [10pts] Let X be a Gaussian random vector with

$$E[X] = \begin{pmatrix} 10 \\ 5 \end{pmatrix}, \quad cov(X) = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

Write down an explicit expression for the p.d.f. of X. i.e. if $X = (x_1, x_2)$, write down the joint p.d.f. $P(x_1, x_2)$

(2) [10pts] Let A, B be the $p \times q$ matrices and x be a random q vector. Prove that

$$cov(Ax, Bx) = Acov(x)B^T,$$

where $cov(u, v) := E[(u - E(u))(v - E(v))^T]$ is the covariance matrix between two random vectors, while $cov(u) := E[(u - E(u))(u - E(u))^T]$ is the covariance matrix for u.

Exercise 3 [20pts]

(1) [10pts] Given the distribution $\frac{1}{3\sqrt{2\pi}}e^{-x^2/18}$, what is the probability that $x > 2$? Express your answer in terms of the error function $\operatorname{erf}(\xi) = \frac{2}{\sqrt{\pi}}\int_0^{\xi} e^{-t^2}\,\mathrm{d}\,t$.

(2) [10pts] For each of the following probability distributions, apply the Markov inequality and Chebyshev inequality to bound $\Pr[x \geq a]$, where $a > 1$. For what value(s) of a is the Chebyshev bound tighter in each case?

a, $p(x) = \begin{cases} 1, & x = 1 \\ 0, & \text{otherwise} \end{cases}$

b, $p(x) = \begin{cases} \frac{1}{2}, & 0 \leq x \leq 2 \\ 0, & \text{otherwise} \end{cases}$

Exercise 4 [20 pts]

Let A and B be $n \times n$ matrices with mutually independent entries sampled from the Poisson distribution with parameter λ. (Hence, $\mathrm{E}[a_{ij}] = \mathrm{Var}(a_{ij}) = \lambda$ and $\mathrm{E}[b_{jk}] = \mathrm{Var}(b_{jk}) = \lambda$.) Compute

$$\mathrm{E}\left[\|AB\|_{\mathrm{F}}^2\right],$$

and express your answer in terms of n and λ. Show your derivations.

Exercise 5 [40pts]

In this exercise, you will first set up a Juypter notebook programming and execution environment and then use it for subsequent exercises. You need to refer the book *Dive into Deep Learning* [PDF].

For part (3) below, you DO NOT need to submit your code. Just attach some screenshots of your running result and write a report

(1) Read code and installation pages 5, 9, 10, 11 of the book. Download the code from the book[link] and install the runtime environment.

(2) (10 pts) For these subproblems, you can refer to Chapter 2 section 2.3 in the book (starting from page 54).

(2.a) (5 pts) The tensor X of shape (2, 3, 4) is defined in this chapter section. What is the output of len(X)?

(2.b) (5 pts) For a tensor X of arbitrary shape, does len(X) always correspond to the length of a certain axis of X? What is that axis?

(3) (30 pts) Select a data set from `https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/` and redesign a MLP on it. Run and analyze the result. You are encouraged to experiment on different configurations. You can refer to Chapter 4 section 4.2 in the book (starting from page 140). Write a report of your findings.