Homework Assignment # 1

Due: Friday, January 24, 2025, 11:59 p.m. Mountain time

Total marks: 100

Question 1. [10 MARKS]

Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be iid multivariate Gaussian random variables, with $\mathbf{X}_i \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, with $\boldsymbol{\mu} \in \mathbb{R}^d$ and $\boldsymbol{\Sigma} \in \mathbb{R}^{d\times d}$ for dimension $d \in \mathbb{N}$. Define sample mean estimator $\bar{\mathbf{X}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{X}_i$. In this question you will reason about its properties, extending beyond the univariate sample mean you reasoned about before.

(a) [1 MARK] What is the dimension of $\mathbb{E}[\bar{\mathbf{X}}]$?

(b) [2 MARKS] What is $\mathbb{E}[\bar{\mathbf{X}}]$? Write it explicitly in terms of the given variables $\boldsymbol{\mu}, \boldsymbol{\Sigma}$ and n. Note that you do not necessarily need to use all three. Show all you steps.

(c) [3 MARKS] What is $\mathrm{Var}[\bar{\mathbf{X}}]$? Write it explicitly in terms of the given variables $\boldsymbol{\mu}, \boldsymbol{\Sigma}$ and n. Note that you do not necessarily need to use all three. Show all you steps.

(d) [2 MARKS] Now define $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{d} X_{ij}$ where X_{ij} is the j-th element in $\mathbf{X}_i$. What is $\mathbb{E}[\bar{Y}]$? Show all you steps.

(e) [2 MARKS] Show that $\mathrm{Var}[\bar{Y}] = \frac{1}{n}\sum_{j=1}^{d}\sum_{k=1}^{d} \Sigma_{jk}$.

Question 2. [15 MARKS]

The purpose of this question is to understand the trade-offs in the different strategies in Chapter 3 to deal with rank-deficient linear systems. Suppose we are given the following dataset

$$\mathbf{X} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{n\times d} \qquad \mathbf{y} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \in \mathbb{R}^{n}$$

where the rows correspond to each sample, with $n = 3$ and $d = 3$. Our goal is to find a set of weights $\mathbf{w} \in \mathbb{R}^d$ such that

$$\mathbf{w} \in \arg\min_{w\in\mathbb{R}^3} \|\mathbf{X}w - \mathbf{y}\|_2^2 \,.$$

(a) [3 MARKS] Is the $\mathbf{w} \in \mathbb{R}^3$ unique? Why or why not?

(b) [3 MARKS] Compute the SVD of $\mathbf{X}$ and use this to provide a closed-form solution for $\mathbf{w}$.

(c) [3 MARKS] Find a $\mathbf{w} \in \mathbb{R}^3$ by performing 2 steps of gradient descent with $\mathbf{w}_0 = (0, 0, 0)$ and $\eta = 0.5$. (Note: you can do these two steps by hand or using code). Now what if $\mathbf{w}_0 = (1, 1, 1)$? Which is better. Compare to the solution given by SVD.

(d) [3 MARKS] Finally try different initializations. What do you notice?

(e) [3 MARKS] Solve the below optimization problem

$$\mathbf{w}_\lambda \in \arg\min_{w\in\mathbb{R}^3} \|\mathbf{X}w - \mathbf{y}\|_2^2 + \lambda\|w\|_2^2$$

for $\lambda \in \{10^{-16}, 10^{-2}, 10^{0}\}$. How do these weights compare to those you found in the above questions?

Question 3. [75 MARKS]

Complete the programming assignment in the associated julia file called `A1.jl`.

Homework policies:

Your assignment should be submitted on eClass as a single pdf document and a zip file containing: the code and a pdf of written answers. The answers must be written legibly and scanned or must be typed (e.g., Latex).

We will not accept late assignments. Plan for this and aim to submit at least a day early. If you know you will have a problem submitting by the deadline, due to a personal issue that arises, please contact the instructor as early as possible to make a plan. If you have an emergency that prevents submission near the deadline, please contact the instructor right away. Retroactive reasons for delays are much harder to deal with in a fair way.

All assignments are individual. All the sources used for the problem solution must be acknowledged, e.g. web sites, books, research papers, personal communication with people, etc. Academic honesty is taken seriously; for detailed information see the University of Alberta Code of Student Behaviour.

Good luck!