Indian Institute of Technology Kharagpur
Class Test 01 2021-22

Date of Examination: **28 Jan. 2022** **Duration:** **40 minutes**
Subject No.: **CS60010** **Subject:** **Deep Learning**
Department/Center/School: **Computer Science** **Credits:** **3** **Full marks:** **20**

Instructions

i. This question paper contains 2 pages and 3 questions. All questions are compulsory. Marks are indicated in parentheses. This question paper has been cross checked.

ii. Please write your name, roll number, subject name and code, date and time of examination on the answer script before attempting any solution.

iii. **Organize your work**, in a reasonably neat and coherent way. Work scattered all across the answer script without a clear ordering will receive very little marks.

iv. **Mysterious or unsupported answers will not receive full marks**. A correct answer, unsupported by calculations, explanation, will receive no marks; an incorrect answer supported by substantially correct calculations and explanations may receive partial marks.

v. In the online mode of the quiz, you need to upload yuor answer scripts as **pdf file**. You can scan your worked out example or you can use latex to produce the pdf.

---

1. (a) (5 points) If $\mathbf{A}$ is $p \times q$ matrix, $\mathbf{U}$ is a $p \times p$ orthogonal matrix and $\mathbf{Z}$ is a $q \times q$ orthogonal matrix, prove that $||\mathbf{A}||_2 = ||\mathbf{UAZ}||_2$.

Solution:

$$||\mathbf{UAZ}||_2 = \max_{\mathbf{x} \neq 0} \frac{||\mathbf{UAZx}||_2}{||\mathbf{x}||_2} \tag{1}$$

Now, length preserving property of orthogonal matrix $\mathbf{U}$ implies $||\mathbf{UAZx}||_2 = ||\mathbf{AZx}||_2$. This, in turn, means [using eqn. (1)].

$$\begin{aligned} ||\mathbf{UAZ}||_2 &= \max_{\mathbf{x} \neq 0} \frac{||\mathbf{AZx}||_2}{||\mathbf{x}||_2} \\ &= \max_{\mathbf{x} \neq 0} \frac{||\mathbf{AZx}||_2}{||\mathbf{Zx}||_2} \quad \text{[Length preserving property of } \mathbf{Z}\text{]} \end{aligned} \tag{2}$$

Now $\mathbf{Z}$ is orthogonal and so is non-singular. Also, $\mathbf{x} \neq 0$ by definition. Hence, $\mathbf{Zx}$ is a non-zero vector, say $\mathbf{y}$. Using $\mathbf{Zx} = \mathbf{y}$ in eqn. (2), we get,

$$||\mathbf{UAZ}||_2 = \max_{\mathbf{y} \neq 0} \frac{||\mathbf{Ay}||_2}{||\mathbf{y}||_2} = ||\mathbf{A}||_2 \quad \text{[Hence proved]}$$

2. (a) (6 points) Prove Euclidean balls are Convex Sets.

Hint: Eucliean balls are represented as $B = \{x \mid ||x - x_0||_2 \leq r\} = \{x \mid (x - x_0)^T (x - x_0) \leq r^2\} = \{x_0 + r\mu \mid ||\mu|| \leq 1\}$.

(b) (2 points) Prove that pointwise maximum operation i.e. $f(x) = \max(f1(x), f2(x))$ preserves convexity.

Solution:
(a) Let x_1 and $x_2 \in B$,

$$\Rightarrow ||x_1 - x_0||_2 \leq r \ \&\ ||x_1 - x_0||_2 \leq r$$

If we can prove $z = \lambda x_1 + (1-\lambda)x_2 \in B$ for any $\lambda \in [0,1]$, B is a convex set.

$$\begin{aligned}
\Rightarrow ||\lambda x_1 + (1-\lambda)x_2 - x_0||_2 &= ||\lambda(x_1 - x_0) + (1-\lambda)(x_2 - x_0)|| \\
&\leq ||\lambda(x_1 - x_0)|| + ||(1-\lambda)(x_2 - x_0)|| \\
&= \lambda||x_1 - x_0|| + (1-\lambda)||x_2 - x_0|| \\
&= \lambda r + (1-\lambda)r \\
&= r
\end{aligned}$$

Hence $||z - x_0|| \leq r \Rightarrow z = \lambda x_1 + (1-\lambda)x_2 \in B$. This implies B is a convex set.
(b) To prove that pointwise maximum preserves convexity, it is enough to show that for any two convex functions f_1, f_2, the function $\max(f_1, f_2)$ is convex. This means the following should be shown,

$$\max(f_1, f_2)(\lambda x + (1-\lambda)y) \leq \lambda \max(f_1, f_2)(x) + (1-\lambda)\max(f_1, f_2)(y)$$

Starting with the LHS,

$$\begin{aligned}
\max(f_1, f_2)(\lambda x + (1-\lambda)y) &= \max(f_1(\lambda x + (1-\lambda)y), f_2(\lambda x + (1-\lambda)y)) \\
&\leq \max(\lambda f_1(x) + (1-\lambda)f_1(y), \lambda f_2(x) + (1-\lambda)f_2(y)) \\
&\leq \lambda \max(f_1, f_2)(x) + (1-\lambda)\max(f_1, f_2)(y)
\end{aligned}$$

Therefore pointwise maximum preserves convexity.

3. (6 points) Let $X_1, X_2, ..., X_n$ be samples from $U(0, \theta)$ or a uniform distribution with parameters $a = 0, b = \theta$. Derive the maximum likelihood estimate for θ using the samples $\{X_i\}_{i=1}^n$.

Solution:

$$X_1, X_2, ..., X_n \sim U(0, \theta)$$

The likelihood function

$$l(\theta|X) = \begin{cases} \left(\frac{1}{\theta}\right)^n, & \theta \geq X_{(n)} \\ 0, & \text{otherwise} \end{cases}$$

where $X_{(n)}$ is the n-th order statistic – in this case, the $\max\{X_1, X_2, ..., X_n\}$ This is a monotonously decreasing function from $X_{(n)}$ to ∞. The maximum value of $l(\theta|X)$ is at $X_{(n)}$ or $\max\{X_1, X_2, ..., X_n\}$.

So, Maximum likelihood estimate is $\max\{X_1, X_2, ..., X_n\}$.