Indian Institute of Technology Kharagpur
Class Test I: 2020-21

Date: 19 Jan. 2021 **Subject No.: CS60010** **Subject: Deep Learning**

2. (a) (3 points) Suppose we have a cost function

$$J(\boldsymbol{\theta}) = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\theta}^T \mathbf{x}^{(i)} + b\mathbf{y}^{(i)} + \frac{1}{2}\boldsymbol{\theta}^T \mathbf{A}\boldsymbol{\theta}$$

where $\boldsymbol{\theta} \in \mathbb{R}^d$ is the parameter vector $\mathbf{x}^{(i)} \in \mathbb{R}^d$, $y^{(i)} \in \mathbb{R}$, $\{\mathbf{x}^{(i)}, y^{(i)}\}$ are N training data points, $\mathbf{A} \in \mathbb{R}^{d\times d}$ is a symmetric matrix and $b \in \mathbb{R}$. We want to find parameters $\boldsymbol{\theta}$ using gradient descent. Find the vector of partial gradients of the cost function.

(b) (1 point) Give the closed-form solution of θ from the above expression you found.

(c) (4 points) Let λ and $\mathbf{x}$ are respectively the eigenvalue and eigenvector of a square matrix $\mathbf{A}$. Prove that $\mathbf{x}$ is also an eigenvector of $\mathbf{A}^k$ where k is a positive integer. Also prove that λ^k is the eigenvalue of $\mathbf{A}^k$.