|  |  |  |  |
| --- | --- | --- | --- |
|  | **PES University, Bengaluru**  (Established under Karnataka Act No. 16 of 2013) | | **UE20CS904** |
| **November 2022 : END SEMESTER ASSESSMENT (ESA)**  **M TECH DATA SCIENCE AND MACHINE LEARNING\_ SEMESTER I**  **UE20CS904 - Mathematical Foundation** | | | |
| Time: 3 Hrs | | Answer All Questions | Max Marks: 80 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Section A (20 marks)** | | | |
| 1 | a) | Find if | 2 |
| sol |  |  |
| b) | Which distance metric is suitable for calculating the least number of squares moved between the starting position (Green Point) and ending position (Blue point) on the chessboard (each square of unit length) for the Queen (Queen can move either diagonally or vertically or horizontally)? Give formula for the same. | 2 |
| sol | Solution: Manhattan distance or Euclidean distance and Chebyshev distance can be used to calculate the distance between two points.  **green = (1,6) blue = (5,3)**  The Manhattan distance or L1 norm is,  **L1 Norm** : ||u|| = |**x1**| + |**y1**| = 4 + 3= 7  The Euclidean distance or L2 norm is,  **L2 Norm** : ||u|| = = = = = 5  **Chebyshev Distance**  **Norm:** ||u||= |  |
| c) | Find out whether the function is concave or convex , | 2 |
| sol |  |  |
| d) | Find the vector projection of the vector a = {5; 5} on b = {8; 2}  Vector projection of a on b =((a . b) / (b . b)) b  = (50/68) (8,2)  = (5.9, 1.5) | 2 |
|  | e) | What is the effect of higher learning rate in Gradient descent algorithm?  When the learning rate is a higher value, it can cause the model to converge too quickly to a suboptimal solution. | 2 |
|  | | | |
| 2 | a) | Calculate the angle between two given vectors. The two vectors are,  a = + 2 and  b = 9 + 3 | 2 |
| sol | Solution :  cos Ɵ = =  = =  = =  Hence Ɵ = 45° |  |
| b) | **What will happen when eigenvalues are roughly equal?**  A. PCA will perform outstandingly  B. PCA will perform badly  C. Can’t Say  D. None of above  Sol: B. PCA will perform badly Since All Pcs are same  When eigenvalues are roughly equal, this indicates that each principal component explains a similar amount of variance in the data. In such cases, dropping any principal components would result in a significant loss of information. This makes PCA less effective for dimensionality reduction because it implies that the original data is uniformly spread across all dimensions and there are no significantly dominant directions of variance to prioritize. | 2 |
|  | c) | Calculate the Jacobian matrix for the following function  𝑓1(x,y) = x3y  𝑓2(x,y) = + y2  Solution : | 2 |
| d) | In simple linear regression for a single data point we define loss as  where is predicted value for find and |
|  | d) |  | 2 |
| e) | We have an rgb image saved as img. An RGB image has length and width 63.We are creating a new image by concatenating img[:,:63,1] , img[:,63:126,:2] & img[:,126:,0]. Wha kind of changes can we observe in the new image as compared to the original image (img).  Solution :  img[row\_start:row\_end, column\_start:column\_end , channels]  red\_channel = img[:,:,0]  **img[:,:63,1]:**  This slices the original image to take all rows (:), the first 63 columns (:63), and only the green channel (1) of the RGB image. This effectively extracts a green-channel-only view of the left third of the original image.  img[:,63:126,:2]:  There seems to be a small confusion in the slicing. Given that the original image width is 63 pixels, img[:,63:126,:2] suggests an attempt to slice beyond the image's width.  img[:,126:,0]:  Similar to the second operation, this suggests taking a slice beyond the described dimensions of the image. | 2 |
| **Section B (30 marks)** | | | |
| 3 | a | Consider a firm operating two plants in two different locations. They both produce the same output (say, **10 units**) using the same type of inputs. Although the amounts of inputs vary between the plants the output level is the same.   1. The firm management suspects that the production cost in Plant 2 is higher than in Plant 1. Verify?   The following information was collected from the managers of these plants. | 5 |
| b | **(ii) A stone is dropped** into a **quiet lake** and waves move in circles at a speed of 5cm per second. At the instant, when the radius of the circular wave is 8 cm, how fast is the enclosed area increasing?  Solution :  we'll use the relationship between the area of a circle and its radius to find out how fast the enclosed area is increasing when the radius of the circular wave is 8 cm. The area A of a circle is given by the formula A=πr2, where r is the radius of the circle. To find how fast the enclosed area is increasing, we need to find the rate of change of the area with respect to time, denoted as dA/dt​.    https://lh4.googleusercontent.com/1Po_iM4DyK_q8BnbMIzj4lTBJkAJ3lQ6P-4LQDUA-UnyL7kD1cIMWx49m3uJ9Fz_69izv8iMzb1FnYkIMHgfvdRE99AeKw-ueQmVcer0cYXSWUHtvJ3UIK2FciCQvw | 5 |
|  | c) | Compute the following convolution for coloured cell, what kind of output the following convolution will have on an image? | 5 |
|  | d) | In the plot shown below the dark shaded portion represents the original coordinates of an object and the same after transformation is represented by the lightly shaded object. Write the coordinates, the transformation matrix and the coordinates after transformation.  https://lh5.googleusercontent.com/xUVrZ26cKgJMvCe7exeIWEneR1cy12gPKg6qhNE2ZcqbQzNZtycN1AhpGZA68f5sZ__NUvNvE-lijNtuXWiN-2oV6CYPkuMU8i02oS5cmbV-UtnYmpRWVKTb2-LI_vf6ArWs32Uh | 5 |
|  | e) | The Following table lists the weight and heights of 5 boys Find the covariance matrix for the data.   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | Boy | 1 | 2 | 3 | 4 | 5 | | Weight(lb) | 120 | 125 | 125 | 135 | 145 | | Height(in.) | 61 | 60 | 64 | 68 | 72 | | 5 |
|  | f) | Find the Eigen values of A : | 5 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Section C (30 marks)** | | | |
| 4 | a) | A headphone manufacturer determines that in order to sell x units of a new headphone,  the price per unit, in dollars, must be p(x) = 1000 - x.  The manufacturer also determines that the total cost of producing x units is given by  C(x)= 3000 + 20x.  i) Find the total revenue R(x)  ii) Find the total profit P(x).  iii) How many units must the company produce and sell in order to maximize profit?  iv) What is the maximum profit?  v) What price per unit must be charged in order to make this maximum profit?  Sol:- | 10 |
| b) | Find singular Value decomposition of | 10 |
|  | c) | Consider the data given below and fit a linear regression line y = ax+b using gradient descent.  X 0 0.4 0.6 1  Y 0 1 0.48 0.95  Initialize the weights a and b to 0.8, 0.2 respectively. Update the weights such that the error is minimum using gradient descent. Use the function sum of squared errors∑ 〖(y-y ̂)〗^2 where y ̂ is the y-predicted value and y is the actual given y. Plot the linear regression line after updating  sol  See the plot of the linear regression line which is close the the new a and b values    See the plot of the linear regression line which is close the the new a and b values | 10 |

then we minimum