Birla Institute of Technology and Science, Pilani

Work Integrated Learning Programmes Division

Cluster Programme - M.Tech. in Data Science and Engg.

II Semester 2020-21

Course Number	DSECL ZC416
Course Name	Mathematical Foundation for Data Science
Nature of Exam	Open Book
Weightage for grading	40%
Duration	120 minutes
Date of Exam	26/09/2021 (10:00 - 12:00)

# Pages	4
# Questions	4

Q1 Basics of linear algebra

a) Given matrices $\mathbf{A}, \mathbf{U}, \Sigma$ and $\mathbf{V}$, write a pseudocode to determine if $\mathbf{U}\Sigma\mathbf{V}^T$ is the SVD of $\mathbf{A}$. You may use the function $[\mathbf{E}, \mathbf{F}] = \text{eigs}(\mathbf{X})$ to determine the eigenvectors $\mathbf{E}$ corresponding to the eigenvalues in the diagonal elements of $\mathbf{F}$, for the square matrix $\mathbf{X}$. Other functions that are needed are to be written. Ensure that everything including the size of the matrices are checked and appropriate error messages are printed. Allocate memory for the data types wherever necessary. Usage of direct multiplication to check if $\mathbf{U}\Sigma\mathbf{V}^T$ is equal to $\mathbf{A}$ should not be done and would not be awarded any marks. (4)

b) Given a matrix $\mathbf{A}$ of size $2m \times m$, with $m > 12$, Prof. Vinod asks his students if in the matrix $R(= r_{ij})$, got through QR decomposition of $\mathbf{A}$, whether $r_{22} > 0$. One student Raj says yes but another student Vinay says no. Who is right and why? In case the question does not have enough data to answer, point out the missing things. (2)

c) Solve the below LPP without using the simplex method (2)

$$\text{Max } Z = x_1$$

subject to

$$a_j x_1 + x_{j+1} \leq b_j \text{ for } j = 1, 2, \ldots, n$$

where $\mathbf{b} \geq 0, \mathbf{a} > 0$ and $\mathbf{x} \geq 0$ are vectors of suitable sizes.

d) Let $T : V \to V$ be a linear transformation from a vector space V onto itself. Suppose $x \in V$ is such that $T^4 x = 0$, $T^3 x \neq 0$. Prove or disprove that x, Tx, $T^2 x$ and $T^3 x$ are linearly independent. (2)

Q2 Fundamentals of optimization and counting

a) The organizing commitee of a cricket match wants to maximize the profit P from the sales of the tickets sold. The cricket stadium will have three types of seating arrangement with their corresponding ticket rates. The maximum seating capacity of the stadium is 15000. The committee wants to allocate the number of seats in the stadium for each category so that profit P is maximum which is defined as $P = 500A^{0.45}B^{0.4}C^{0.2}$, where A, B and C are the number of seats in each category. It is decided that the ticket rates in category A, B and C are Rs. 550, Rs. 1800 and Rs. 3000 respectively.

   i) What are the different constraints that needs to be applied to solve the problem ?

   2) Find the optimum number of seats in each category by optimizing the profit subject to above constraints, and the maximum profit. (1+2)

b) A researcher needs to minimize a function while working on his research problem given by

$$f(x_1, x_2) = 9x_1^2 + x_2^2 + 18x_1 - 4x_2$$

If he selects the steepest descent algorithm to minimize the function, what will be the values of x_1, x_2, α_k and f in the first five iterations of the algorithm? Here, α_k is the step size in the k^{th} iteration. You may take the starting point as $(-2, -3)$. (3)

c) A software engineer needs to find the critical points of a multivariate function $g(x, y) = x^4 + y^3 - 5x^2y - 2xy^2 + x^2$. Construct a table to classify the critical points of the given function as local minima, local maxima or saddle point. (2)

d) Give a recursive definition of the function, which counts the number of ones in a bit string s. (1)

e) Use structural induction to show that $l(T)$, the number of leaves of a full binary tree T, is 1 more than $i(T)$, the number of internal vertices of T. Show all the steps. (1)

Q3 Fundamentals of induction and combinatorics

a) Pradheep feels that the he cannot prove the inequality given by

$$\frac{1}{2}.\frac{3}{4}\cdots\frac{2n-1}{2n} < \frac{1}{\sqrt{3n}}$$

for all positive integers n using mathematical induction. Is his claim correct? (1.5)

b) Pradeep's teacher Ramesh, looking at the inequality in a) part above suggests that he could try proving a stonger inequality

$$\frac{1}{2}.\frac{3}{4}\cdots\frac{2n-1}{2n} < \frac{1}{\sqrt{3n+1}}, \quad n \geq 1$$

using mathematical induction and use an argument to prove part a). Is Ramesh correct in suggesting this? If so, how to prove a) using b)? If not, justify. (1.5)

c) A two dimensional function f, whose domain in $\mathcal{N} \times \mathcal{N}$ is defined by

$$f(a,b) = \begin{cases} 2b, & \text{if } a = 0 \\ 0, & \text{if } a \geq 1 \text{ and } \ b = 0 \\ 2, & \text{if } a \geq 1 \text{ and } \ b = 1 \\ f(a-1, f(a, b-1)), & \text{if } a \geq 1 \text{ and } \ b \geq 2 \end{cases}$$

i) Can $f(a,2) = 4$ for $a \geq 1$ be established using induction? (1)

ii) Evaluate $f(2,3)$ and $f(3,3)$? (1)

d) In National Games Championship, a high jump competition is being conducted along with other athletic events. Only n out of m athletes were able to meet the eligibility criteria set by the athletics committee of the games. In this event, multiple athletes can even end up jumping the same height and hence, the awards committee has even permitted tie among the athletes as a valid result. Under these assumptions how many different results are possible at the end of competition. (2)

e) A programming language was designed in such a way that all variable names are of length exactly six and is built out of only lowercase alphabets from the English language. The compiler designer wants to know the maximum number of variable names such that

i) the letter h appears in the variable name?

ii) both letters h and i appear in the variable name?

iii) the letters h and i appear side by side with the condition that h appears first. In this case assume that all six letters are distinct in the variable name. (1+1+1)

Q4 Advanced counting, recurrence relations and generating functions

a) Let n be a non-negative integer. Consider the functions $f(n) = \frac{(2n)!}{(2n-2)!2!}$, $g_2(n) = \frac{n!}{2!(n-2)!} + n \times n$ and $g_1(n) = \frac{n!}{2!(n-2)!}$

   i) Can you prove that $f(n) = g_1(n) + g_2(n)$ by using a combinatorial counting approach by showing that LHS and RHS are counting the same object?

   ii) Can you prove that $f(n) = g_1(n) + g_2(n)$ by using an algebraic approach? (2+1)

b) Using generating functions or otherwise, solve the recurrence relation defined by $H_k = 3H_{k-1} + 4^{k-1}$, for $k \geq 1$, with the initial condition $H_0 = 1$. Is the solution unique? Justify. (3)

c) A sequence $\{f_n\}_{n \in \mathbb{Z}^+}$, is defined by $f_n = 2f_{n-1} + g(n)$, where $g(n) = 2^n$. Can you prove that $r_n = ng(n)$ is a solution? Also, if possible, derive the the general solution by assuming a suitable value for f_0. (2)

d) A gasoline car dealer is stopping operations in a city as part of reallocation of company resources to electric cars. It has eight identical sedan cars left in its inventory. They give an advertisement in a local newspaper announcing free distribution of remaining cars as a symbol of gratitude to their city of operation. But only five customers show up on the final day. Can you find out the total number of ways by which one can distribute these cars to these five customers such that a given customer does not receive more than two cars but should receive at least one car. (2)