COL 341: Fundamentals of Machine Learning

22/11/2019 **Major Examination** Maximum Marks: 80

[20 Marks]

Q1. For each of the following machine learning methods, write the corresponding optimization problems that need to be solved. Write the precise mathematical formula for the objective function, whether it is a minimization or a maximization (or any other type) problem, the number and types of variables over which optimization is being done and constraints (if any), assuming that there are n-input examples of the form $x_i \in R^m$ for $i \in [1..n]$ and the corresponding labels are y_i. You don't need to solve the problem, or give any justification, you just need to write the corresponding optimization problem precisely. If you need to use any additional symbols and notations, define them before you use.

(i) Linear regression (ii) Ridge regression (iii) Lasso regression (iv) Logistic regression (v) Support vector machine (vi) K-means (vii) Principal component analysis (viii) Non-negative matrix factorization with missing values in x_i's (ix) Generative adversarial networks (x) Single node decision tree (write down the objective that is being optimized while determining the split at the root node)

[20 Marks]

Q2. Let X be a random vector of size m, distributed according to multivariate Gaussian distribution with a mean $\mu \in R^m$, and covariance matrix $\Sigma \in R^{m \times m}$. Let A be a n x m matrix. We will try to find the distribution of a vector of n random variables given by Y = AX.

(i) Find the mean of the random variables Y.

(ii) Show that the covariance matrix of Y is $A \Sigma A^T$

(iii) **[Extra credit]** Suppose we wish to find a transformation matrix A such that the resulting variables Y = AX are independent of each other. How would you find such A (hint, use SVD/PCA and the fact that Σ is symmetric)?

[20 Marks]

Q3. Let A be a n x m matrix with the following Singular Value Decomposition $A = U D V^T$. We wish to approximately represent the columns of the matrix $B = A^TA$ using a k-dimensional subspace (k << m) such that $B_{.,j} = \sum_{i=1}^{k} \alpha_{ij} z_i$, where $z_i \in R^m$ are the k basis vectors of the k-dimensional subspace of R^m. Find a set of vectors z_i's and the corresponding low dimensional representation of B that minimize the error (you can use the well-known results from PCA/SVD to get your answer).

[20 Marks]

Q4. Consider a hypothesis class of axis-aligned rectangles in 2-dimensional space. An axis-aligned rectangle is represented by two points: its bottom left corner (x1, y1) and its top right corner (x2, y2) (x1 < x2 and y1 < y2). Any point (x, y) is classified as positive if and only if it lies inside the rectangle, i.e., x1 ≤ x ≤ x2 and y1 ≤ y ≤ y2. In this question, you will find the VC dimension of this hypothesis class.

(i) Suitably position three points in 2-dimensional space and show that the axis aligned rectangles can shatter your points.

(ii) Show that no set of 5 points in 2-dimensional space can be shattered by the axis aligned rectangles. Hint: Consider the following four points from any set of five points and assign them class labels suitably: point with smallest x-coordinate, point with largest x-coordinate, point with smallest y coordinate, point with largest y coordinate.

(iii) Can axis aligned rectangle shatter 4 points in 2-dimensional space? Prove or disprove. What is the VC dimension of hypothesis class of axis-aligned rectangles?