Homework #0 B

Spring 2020, CSE 446/546: Machine Learning
Dino Bektesevic

Collaborated: Conor Sayers, Joachim Moeyenes, Jessica Birky, Leah Fulmer

Probability and Statistics

B.1 *[1 points]* Let $X_1, \ldots, X_n$ be n independent and identically distributed random variables drawn uniformly at random from $[0, 1]$. If $Y = \max\{X_1, \ldots, X_n\}$ then find $\mathbb{E}[Y]$.

Linear Algebra and Vector Calculus

B.2 *[1 points]* The *trace* of a matrix is the sum of the diagonal entries; $Tr(A) = \sum_i A_{ii}$. If $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{m\times n}$, show that $Tr(AB) = Tr(BA)$.
It can be shown in general that trace is invariant under cyclic permutations, which for case of n=2 looks like commutation:

$$\mathrm{tr}(AB) = \sum_i^m (ab)_{ii} = \sum_{i=1}^m \sum_{j=1}^n a_{ij} b_{ji} = \sum_{j=1}^n \sum_{i=1}^m b_{ji} a_{ij} = \sum_i^n (ab)_{jj} = \mathrm{tr}(BA)$$

using the fact that product of $n \times m$ and $m \times n$ matrix is an $m \times m$ matrix and that the trace of a square matrix can be rewritten as:

$$\mathrm{tr}(A^T B) = \sum_{i,j} A_{ij} B_{ij}$$

as per Wikipedia.

B.3 *[1 points]* Let $v_1, \ldots, v_n$ be a set of non-zero vectors in $\mathbb{R}^d$. Let $V = [v_1, \ldots, v_n]$ be the vectors concatenated.

a. What is the minimum and maximum rank of $\sum_{i=1}^n v_i v_i^T$?
Vector v is a column vector with n rows. Transpose of vector v will have n columns and 1 row. The matrix resulting from matrix multiplication has the number of rows of the first and the number of columns of the second matrix, i.e. it will be an (n, n) square matrix. Basis of an $\mathbb{R}^d$ is spanned by d linearly independent vectors. So the maximum rank of a matrix will be d when $n \geq d$. Otherwise the maximal rank of the sum is n when $0 < n < d$. If all vectors are the same unit vectors then the minimum rank could be 1.

b. What is the minimum and maximum rank of V?
Same as above with the additional comment that minimal rank of a matrix would be 0 in case of a zero matrix, but since the vectors v must be non-zero the minimal rank of V would again be 1 in the case all v are unit vectors, or if all given vectors are linearly dependent.

c. Let $A \in \mathbb{R}^{D\times d}$ for $D > d$. What is the minimum and maximum rank of $\sum_{i=1}^n (Av_i)(Av_i)^T$?

d. What is the minimum and maximum rank of AV? What if V is rank d?