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 $X_1, \dots, X_n$ be n independent and identically distributed random variables drawn uniformly at random from $[0, 1]$. If $Y = \max\{X_1, \dots, X_n\}$ 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 $v_1, \dots, v_n$ be a set of non-zero vectors in $\mathbb{R}^d$. Let $V = [v_1, \dots, v_n]$ 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 could 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 $\sum_{i=1}^n (Av_i)(Av_i)^T$?
Rank can not be greater than either number of rows or columns, i.e. $\mathrm{Rank}(A) \leq \min(i, j)$. So the rank will be limited to d despite $D > d$. Rank will be preserved is A is monomorphism, i.e. if the operator represented by A is injective. Assuming an monomorphic map we can repeat everything stated above - the rank of V is set between 1 and d depending on the choice of v_i vectors. In any other situation the rank of A can only be less than d as it must project different values from its domain into the same value in its codomain.

d. What is the minimum and maximum rank of AV? What if V is rank d?
The number of columns in the first matrix must be equal to the number of rows in the second matrix, so the product of AV results in an $(D, d)x(d, n) = (D, n)$ matrix. Following what was said above the rank will be limited to the smaller of the two values D or n. The maximum rank of the product AV will be limited by the choice of v's and the properties of mapping A represents.