CSE546 Assignment 0
B Problems

Joseph David

April 5 2021

1 Probability and Statistics

B.1 First, note that since the X_i are i.i.d, for $x \in [0,1]$

$$P(Y \leq x) = \prod_{i=1}^{n} P(X_i \leq x) = x^n,$$

where we have used the fact that $P(X_i \leq x) = x$ since they are uniformly distributed on $[0,1]$. Thus,

$$E[Y] = \int_0^1 P(Y > x)\, dx = \int_0^1 1 - x^n\, dx = \frac{n}{n+1}.$$

B.2 Recall Markov's inequality states that for a random variable X and number a,

$$P(X \geq a) \leq \frac{E[X]}{a}.$$

Therefore, for any $x \geq 0$

$$P(X \geq \mu + \sigma x) = P((X-\mu)^2 \geq \sigma^2 x^2) \leq \frac{E[(X-\mu)^2]}{\sigma^2 x^2} = \frac{\sigma^2}{\sigma^2 x^2} = 1/x^2.$$

2 Linear Algebra and Vector Calculus

B.3 For any $i \in \{1, \ldots, n\}$, we have by definition of matrix multiplication that

$$(AB)_{ii} = \sum_{j=1}^{n} A_{ij} B_{ji}.$$

Therefore,

$$\begin{aligned}
Tr(AB) &= \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}B_{ji} \text{ by definition of trace} \\
&= \sum_{j=1}^{n}\sum_{i=1}^{n} B_{ji}A_{ij} \text{ swap order of summation} \\
&= \sum_{j=1}^{n} (BA)_{jj} \text{ matrix multiplication defn} \\
&= Tr(BA) \text{ by definition of trace}
\end{aligned}$$

B.4(a) We see that the diagonal terms of $v_i^T v_i$ must be nonnegative for all i because they are squares of terms in v_i, and since the v_i are nonzero, there must be at least one positive term. Since this is true for all v_i,

$\sum_{i=1}^n v_i^T v_i$ must also have at least one positive term and hence cannot have rank 0. If we set $v_i = [1\,0\,\cdots\,0]$ for all $i = 1, \ldots, n$, we see that $\sum_{i=1}^n v_i^T v_i$ has rank 1, so this must be the minimum.

For the maximum, note that for arbitrary i, $v_i^T v_i$ has row span of at most 1 since each row is a linear combination of v_i when the corresponding element of v_i is nonzero, and otherwise that row will be zero. Therefore, since we have n vectors, $\sum_{i=1}^n v_i^T v_i$ has rank of at most n when $n < d$. However, the rank is also bounded above by d since it is a $d \times d$ matrix. Thus, the rank is bounded above by $\min\{d, n\}$. If $n < d$, then take the v_i's to be the first n standard basis vectors and $\sum_{i=1}^n v_i^T v_i$ will have rank n. If $n \geq d$, take the first d vectors to be equal to the d standard basis vectors and $\sum_{i=1}^n v_i^T v_i$ will have rank d. This shows that the maximum rank of $\sum_{i=1}^n v_i^T v_i$ is in fact $\min\{n, d\}$.

B.4(b) Since each v_i is nonzero, V must have rank at least 1, and choosing all v_i to be equal to the same nonzero vector gives a rank 1 matrix so this is indeed the minimal rank. Since V is a $d \times n$ matrix, its rank is bounded above by $\min\{n, d\}$. Choosing v_i to be equal to the ith standard basis vector, and v_i equal to any nonzero vector when $i > d$, we get that V is has rank equal to $\min\{n, d\}$ so this is the maximal rank.

B.4(c) If A is the zero matrix, then $\sum_{i=1}^n (Av_i)^T (Av_i)$ will also be the zero matrix and hence the minimal rank is 0. Similarly to before we set the v_i equal to the ith standard basis vector when $i \leq d$ and the remaining v_i may be anything else. If $D \geq n$, then we may set the upper right block to the $n \times n$ identity matrix, in which case $\sum_{i=1}^n (Av_i)^T (Av_i)$ has rank n. If $D < n$, we let A be the identity in the upper left block and $\sum_{i=1}^n (Av_i)^T (Av_i)$ will have rank D. Thus, the maximal rank is $\min\{n, d\}$.

B.4(d) Taking A to be the zero matrix, AV is also the zero matrix and hence the minimal rank is 0. From part (b), we know that V has maximal rank equal to $N = \min\{n, d\}$. Thus, the image of V has dimension at most N, which implies that AV's rank is bounded above by N. We claim that rank N may be achieved. Letting $w_1, \ldots, w_N$ be a basis for the image of V. Let A be the matrix that corresponds to the linear transformation defined by appended $D - d$ zeros to the end of each w_i, which is possible since $N \leq d < D$. Then, the image of AV will have dimension of N, because the Aw_i will retain their linear independence. This proves that $N = \min\{n, d\}$ is also the maximal rank for AV.