CS973: Machine Learning for Cyber Security

Quiz III (10 Nov 2024)

Q3. Router Rank **(3 marks)**

It takes a router 25 clock cycles to process a single network packet. Assume that the CPU for this router does nothing but process network packets. Melbo has collected data on daily packets received by the router and total daily CPU usage for an entire month in the form of a $\mathbb{R}^{30\times 2}$ matrix called D where the first column stores number of packets received in each day and the second column stores the daily CPU usage in terms of clock cycles. See for example, the matrix on the right What is the rank of the matrix D?

$$D = \begin{bmatrix} 2 & 50 \\ 15 & 375 \\ \cdots & \cdots \\ 62 & 1550 \\ 42 & 1050 \end{bmatrix}$$

( ) 0

(✓) 1

( ) 2

( ) 30

Justify your answer briefly.

The second column is just the first column multiplied by 25 i.e. if take the first column as the vector $\mathbf{p} \in \mathbb{R}^{30}$ and $\mathbf{q} = [1,25] \in \mathbb{R}^2$ then $D = \mathbf{p}\mathbf{q}^\top$ which means D is at most rank 1. Note that D is not rank 0 since it is not the all-zeros matrix.

Q4. Rank Escalation **(4 marks)**

Melbo has a $d \times d$ matrix A of rank r. Melbi has another $d \times d$ matrix B of rank s. Which all of the following statements are always true (you may select multiple options)?

[✓] We must always have $\text{rank}(A + B) \leq d$

[ ] We must always have $\text{rank}(A + B) = r + s$

[✓] We must always have $\text{rank}(A + B) \leq r + s$

[ ] We must always have $\text{rank}(A + B) \geq r + s$

Justify your answer briefly i.e. why did you select/not-select each option.

A matrix cannot have rank exceed its number of rows or its number of columns hence the first option is always true. The second and fourth options are not always true since we could have $A = I_d$ as the identity matrix that is full rank and $B = -I_d$ i.e. $r = d = s$ but $A + B$ is the all-zeros matrix which is at most rank 1 (or rank 0 by some conventions). The third option is always correct by using the SVD theorem. If $A = \sum_{i \in [r]} \sigma_i \mathbf{u}_i \mathbf{v}_i^\top$ and $B = \sum_{j \in [s]} \lambda_j \mathbf{p}_j \mathbf{q}_j^\top$, then $A + B$ can be written as the sum of at most $r + s$ singular components.

Q5. Missing Matrix **(7 marks)**

Melbu created a 3×3 matrix X but spilled some coffee on the sheet of paper on which the matrix was written out and some entries got erased. However, Melbu recalls that the matrix is rank 1 i.e., it is of the form $X = \mathbf{p}\mathbf{q}^\top$ where $\mathbf{p}, \mathbf{q}$ are two 4-dimensional vectors. Moreover, it is known that the dot product of these two vectors is $\mathbf{p}^\top \mathbf{q} = 103$ and $\mathbf{p}^\top \mathbf{1} = 21$ where $\mathbf{1}$ is the all-ones vector. Help Melbu discover all the missing entries. You may assume that $\mathbf{p}, \mathbf{q}$ contain only integer entries.

$$\begin{bmatrix} 32 & 12 & a & 20 \\ 72 & 27 & b & 45 \\ o & q & r & s \\ 48 & 18 & t & 30 \end{bmatrix}$$

We need to discover $\mathbf{p}, \mathbf{q}$ to solve the problem. The entries of the first column (or any other column for that matter) tell us that the vector $\mathbf{p}$ must be of the form $c' \cdot \left[1, \frac{72}{32}, \frac{o}{32}, \frac{48}{32}\right] = c \cdot \left[4, 9, \frac{o}{8}, 6\right]$ for some constant c. The entries of the any row tell us that the vector $\mathbf{q}$ must be of the form $d \cdot \left[8, 3, \frac{a}{4}, 5\right]$ for some constant d. Now, with this, the $(1,1)$-th entry of the matrix is $cd \cdot 32 = 32$ which suggests that $cd = 1$. Indeed, using $c = 1, d = 1$ recovers all the known entries of the matrix and this is the only integral solution that works. We are told that the sum of entries in the vector $\mathbf{p}$ is 21 which tells us that $\frac{o}{8} = 2$ i.e. $o = 16$ and $\mathbf{p} = [4,9,2,6]$. We are told that the trace of the matrix is 103 i.e., $r = 14$. This means that $\frac{a}{4} = \frac{14}{2}$ i.e. $a = 28$ and $\mathbf{q} = [8,3,7,5]$. This allows us to fill in the rest of the entries.

$$a = 28, b = 63, o = 16, q = 6, r = 14, s = 10, t = 42$$

Q6. Covariances can vary **(6 marks)**

X, Y are two real-valued random variables (not necessarily independent) such that $\text{Cov}(X, Y) = \sqrt{\mathbb{V}[X] \cdot \mathbb{V}[Y]}$. Also let us define the random variables $Z = X + Y$ and $W = X - Y$. Then which one of the following statements must always be true (you may select multiple options)?

[✓] If X, Y are independent, then at least one of them is a constant random variable.

[ ] If X, Y are not independent, then their covariance must necessarily be zero.

[✓] Irrespective of whether X, Y are independent or not, their covariance can never be negative.

[ ] It must always be the case that Z is a constant random variable.

[ ] It must always be the case that W is a constant random variable.

[ ] It must always be the case that $X \equiv c \cdot Y$ for some real-valued constant c.

The first option is correct since if X, Y are independent then $\text{Cov}(X, Y) = 0$ which means $\mathbb{V}[X] \cdot \mathbb{V}[Y] = 0$ which means that either $\mathbb{V}[X] = 0$ or $\mathbb{V}[Y] = 0$ which means at least one of the random variables is a constant (no variance).

The second option is not correct – suppose X is a standard Rademacher random variable i.e. takes value ± 1 with equal probability and define $Y \equiv 2X + 1$ i.e., on any event, Y always takes a value equal to one greater than twice the value taken by X on that event. Clearly these random variables are non-independent. In this case, $\mathbb{V}[X] = 1, \mathbb{V}[Y] = 4$ and $\text{Cov}(X, Y) = 2$ i.e. we do get $\text{Cov}(X, Y) = \sqrt{\mathbb{V}[X] \cdot \mathbb{V}[Y]}$ but $\text{Cov}(X, Y) = 2$ i.e. non-zero.

The third option is correct – if $\text{Cov}(X, Y) = \sqrt{\mathbb{V}[X] \cdot \mathbb{V}[Y]}$ then it can never take a negative value since it is the square root of the product of two variances.

The fourth option is not correct, and the example given above acts as a counter example here as well since Z takes values in the set $\{-1,3\}$ i.e., not constant.

The fifth option is not correct, and the example given above acts as a counter example here as well since W takes values in the set $\{-2,0\}$ i.e., not constant.

The fifth option is not correct, and the example given above acts as a counter example here as well since in the example, Y cannot be obtained by multiplying X by a constant.