Question 4

Ian Dover

February 2023

1 Part A

X is a tensor of size I x J x K.
X can then be reshaped on the nth-mode into the shapes of I x JK, J x IK, or K x IJ. Since the rank of a matrix is determined by the number of linearly independent rows and columns, we can now say that the reshaped representations are less than and equal to JK, IK, or IJ, thus resulting in the inequality $rank(X) \leq min(JK, IK, IJ)$.

To further explain this, imagine you have a matrix with either (a) a single row of non-zero values, or (b) a single column of non-zero values. This matrix can then be represented as a single basis of the outer product of two vectors.

We can then extend this logic to an entire matrix of non-zero values. If we loop through each column or row of the matrix, we can construct a single basis function for each row or column of the matrix. This means that the number of basis used in the construction of the matrix is the minimum of either the number of rows or number of columns. This is also the definition of the rank of the matrix.

2 Part B

2.1 Section 1

We are trying to minimize the problem $||X - ||G; A^{(1)}, A^{(2)}, ..., A^{N}||||^2$.
This problem can be shown in the vector representation (1.1):

$$||vec(X) - (A^{(N)} \bigotimes A^{(N-1)} \bigotimes ... \bigotimes A^{(1)})vec(G)||$$

Considering that all of the factor matrices are orthonormal (1.2):

$$A^T(A \bigotimes X) = X$$

The value for G which minimizes the expression is (1.3):

$$G = X \times_1 A^{(1)T} \times_2 A^{(2)T} ... \times_N A^{(N)T}$$

We can hereby extend this reasoning to the above expression (1.1) and reduce it to (1.4):

$$||vec(X) - vec(X)|| = 0$$

To explain further, by showing that the orthonormality of all factor matrices causes the kronecker-product part of the expression to cancel out the transpose of the tensor dot product, the second term of the expression reduces to $vec(X)$, leading (1.1) to be equivalent to 0.

2.2 Section 2

We need to prove that maximizing the expression:

$$||X \times_1 A^{1^T} \times_2 ... \times_n A^{(n)^T}||$$

Minimizes the expression:

$$||X - ||G; A^1, A^2, ..., A^N||||^2$$

We can rewrite the minimization expression:

$$||X - ||G; A^1, A^2, ..., A^N||||^2$$

$= ||X||^2 - 2 < X, ||G; A^{(1)}, A^{(2)}, ..., A^{(N)}|| > + ||||G; A^{(1)}, A^{(2)}, ..., A^{(N)}||||^2$
$= ||X||^2 - 2 < X \times_1 A^{(1)T} ... x_N A^{(N)T}, G > + ||G||^2$
$= ||x||^2 - ||G||^2$
$= ||x|| - ||X \times_1 A^{1^T} \times_2 ... \times_n A^{(n)^T}||$

Based on the rewrite of the minimization expression, it can be shown based on the negative term that the minimization problem can be solved by the maximization problem:

$$||X \times_1 A^{1^T} \times_2 ... \times_n A^{(n)^T}||$$