VMLS - SG Chapter 10 - Matrix Multiplication Name:__ Ben Chen ________________

Always include this title page with your PDF. Include your name above.

- Submit your work in Gradescope as a PDF - you will identify where your “questions are.”
- Identify the question number as you submit. Since we grade "blind" if the questions are NOT identified, the work WILL NOT BE GRADED and a 0 will be recorded. Always leave enough time to identify the questions when submitting.
- One section per page (if a page or less) - We prefer to grade the main solution in a single page, extra work can be included on the following page.
- Long instructions may be removed to fit on a single page.
- **Do not start a new question in the middle of a page.**
- Solutions to book questions are provided for reference.
- You may NOT submit given solutions - this includes minor modifications - as your own.
- Solutions that do not show individual engagement with the solutions will be marked as no credit and can be considered a violation of honor code.
- If you use the given solutions you must reference or explain how you used them, in particular...

For full credit, EACH book exercise in the Study Guides must use one or more of the following methods and FOR EACH QUESTION. Identify the number the method by number to ensure full credit.

Method 1 - Provide original examples which demonstrate the ideas of the exercise in addition to your solution.

Method 2 - Include and discuss the specific topics needed from the chapter and how they relate to the question.

Method 3 - Include original Python code, of reasonable length (as screenshot r text) to show how the topic or concept was explored.

Method 4 - Expand the given solution in a significant way, with additional steps and comments. All steps are justified. This is a good method for a proof for which you are only given a basic outline.

Method 5 - Attempt the exercise without looking at the solution and then the solution is used to check work. Words are used to describe the results.

Method 6 - Provide an analysis of the strategies used to understand the exercise, describing in detail what was challenging, who helped you or what resources were used. The process of understanding is described.

Method #5

The instructions were not clear… am I supposed to show numerical examples for each or just write them down?

1. (40 pts) Consider pages 179 - 181. Demonstrate with simple matrix examples:

- The 4 properties of matrix multiplication.
- The inner product and matrix-vector products
- A block matrix multiplication
- A column interpretation of matrix-matrix product
- Multiple sets of linear equations (skip row interpretation if you wish)
- Inner product representation
- Gram Matrix
- Outer Product representation

You can use the same simple matrices where possible - however, don't use the identity matrix. You may wish to do more examples to illustrate the additional types of products given. You may use multiple pages.

Associative Property: A(BC) = (AB)C
Distributive Property: A(B+C) = AB + AC and (A+B)C = AC + BC
Multiplicative Identity Property: AI = IA = A
Non-Commutative Property: AB =/= BA

Inner Product numerical example

$u = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} \quad v = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} \quad u^T v = 1(2) + 3(4) + 5(6) = 2 + 12 + 30 = 44$

Matrix-vector Product numerical example

$A = \begin{bmatrix} 1 & 0 & 2 \\ 4 & 1 & 3 \end{bmatrix}, \quad u = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$

$Au \begin{bmatrix} 1(1) + 0(3) + 2(5) \\ 4(1) + 1(3) + 3(5) \end{bmatrix} = \begin{bmatrix} 1+0+10 \\ 4+3+15 \end{bmatrix} = \begin{bmatrix} 11 \\ 22 \end{bmatrix}$

Block matmul

$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$

$AB = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}$

Column Interpretation of Matrix-Matrix Product

$A = \begin{bmatrix} | & | & | \\ a_1 & a_2 & a_3 \\ | & | & | \end{bmatrix} \quad b = [b_1, b_2, b_3] \quad AB = [Ab_1, Ab_2, Ab_3]$

Multiple sets of liner equation

$Ax = b \quad AX = B \quad A \in R^{m \times n} \quad X \in R^{n \times k} \quad B \in R^{m \times k}$

col. are diff — col. multiple right-hand sides

For example

$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad Ax^{(1)} = b^{(1)}$
$Ax^{(2)} = b^{(2)}$

$X = [x^{(1)} \ x^{(2)}] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$Ae_1 = A\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$
$Ae_2 = A\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$

$B = [A(1,0)^T, A(0,1)^T] = [Ae_1, Ae_2]$ so $B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

Inner Product Representation

$u^T v = \sum_{i=1}^{n} u_i v_i \quad u = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad v = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$

$u^T v = [1\ 2]\begin{bmatrix} 3 \\ 4 \end{bmatrix} = 1(3) + 2(4) = 11$

$u^T v = v^T u \quad u^T v = \|u\| \cdot \|v\| \cdot \cos(\theta)$ ← cosine similarity

θ: angle between u, v

Gram Matrix

$G = A^T A$ if $A \in R^{m \times n}$ with col. as $a_1, a_2, \ldots, a_n \in R^m$
$A = [a_1 \ a_2 \ \ldots \ a_n]$

Symmetric Matrix

$G_{ij} = a_i^T a_j$, G is all pairwise inner products of A

is also semi-definite which means all eigenvalues ≥ 0

Outer product Representation

If $u \in R^m$, $v \in R^n$ outer product $= uv^T \quad (uv^T)_{ij} = u_i v_j$

creates rank-1 matrix, express matrices as sum of outer products, low-rank approximations

Method #5

2. (20 pts) Solve and Explain the solution to 10.39 here in your own words. (Since you are given a solution, you will be graded on your ability to explain).

10.39 *Gram matrix and QR factorization.* Suppose the matrix A has linearly independent columns and QR factorization $A = QR$. What is the relationship between the Gram matrix of A and the Gram matrix of R? What can you say about the angles between the columns of A and the angles between the columns of R?

Given A = QR where Q has orthonormal columns (Q transpose Q = I) and R is upper triangular...

the Gram matrix of A is A transpose A

if we take A = QR and substitute it in, A transpose A = (QR) transpose (QR) = R transpose Q transpose QR = R transpose R

A transpose A = R transpose R (ATA = RTR)

We can interpret this as ATA has information about inner products between the columns of A

RTR contains the same information as ATA about the inner products, but is calculated after the columns are presented in the orthonormal basis given by Q, which means that the angles between the columns of A and the angels between the columns of R must be the same (given that they're containing the same info) and so the Gram matrices are also the same.

Gram matrix can calculate the cosines of the angles through:

cos(θij) = (ATA)ij / (||ai|| ||aj||) and same thing for R

Method #5

3. (20pts) Solve and Explain the solution to 10.40 here in your own words. (Since you are given a solution, you will be graded on your ability to explain).

10.40 *QR factorization of first i columns of A.* Suppose the $n \times k$ matrix A has QR factorization $A = QR$. We define the $n \times i$ matrices

$$A_i = \begin{bmatrix} a_1 & \cdots & a_i \end{bmatrix}, \qquad Q_i = \begin{bmatrix} q_1 & \cdots & q_i \end{bmatrix},$$

for $i = 1, \ldots, k$. Define the $i \times i$ matrix R_i as the submatrix of R containing its first i rows and columns, for $i = 1, \ldots, k$. Using index range notation, we have

$$A_i = A_{1:n,1:i}, \quad Q_i = A_{1:n,1:i}, \quad R_i = R_{1:i,1:i}.$$

Show that $A_i = Q_i R_i$ is the QR factorization of A_i. This means that when you compute the QR factorization of A, you are also computing the QR factorization of all submatrices $A_1, \ldots, A_k$.

When you compute the QR factoriazation of the full matrix A, you are simultaneously computing the QR factorization of all the submatrices as well, each Ai has its own QR factorization, where Qi and Ri are just submatrices of Q and R

Given: $A = QR$

$A \in \mathbb{R}^{n \times k}$
$Q \in \mathbb{R}^{n \times k}$ w/ orthonormal cols
$R \in \mathbb{R}^{k \times k}$ is upper triangular

hypothesis: QR factorization is $A_i = Q_i R_i$

$$A_i = A_{1:n,1:i} = Q_{1:n,1:k} R_{1:k,1:i}$$

R is upper triangular so $R_{j,i} = 0$ for $j > i$

$$\rightarrow R_{1:k,1:i} = \begin{bmatrix} R_i \\ 0 \end{bmatrix}$$

only cares about the first i rows

so $A_i = Q_i R_i$

Method #5

4. (20pts) Solve and Explain the solution to 10.43 here in your own words. (Since you are given a solution, you will be graded on your ability to explain).

(If you watched the author videos, this where to take a moment to sit in awe of the speed of computers)

10.43 A particular computer takes about 0.2 seconds to multiply two 1500×1500 matrices. About how long would you guess the computer takes to multiply two 3000×3000 matrices? Give your prediction (*i.e.*, the time in seconds), and your (very brief) reasoning.

matmul has about the runtime complexity of O(n^3) for dense matrices.

So you can calculate it:Time ratio = (3000/1500)^3 = 2^3 = 81500x1500 takes 0.2 seconds and 8 x 0.2 takes 1.6 seconds

But if we're talking about like a sparse matrix, then the matmul operation's multiplication part takes O(n^2) approximately, depending on the sparsity pattern and so the actual complexity ends up being closer to something like O(n*nnz) where nnz = number of non-zero cells/elements in the matrix.

Sparse might be like 0.2 seconds * 4 = 0.8 seconds, but sparsity might need to introduce a search algorithm or something so might vary.