College of Engineering & Applied Sciences

CSPB 2820

Linear Algebra With Computer Science Applications

Study Guide 10 - Matrix Multiplication

Taylor Larrechea

2024

Study Guide 10

Matrix Multiplication

Study Guide Instructions

- 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...

Method Selection

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 or 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.

Problem 1

Problem Statement

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.

Solution

For this problem, the two matrices that we are going to use for demonstration is

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad , \quad B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \quad , \quad C = \begin{bmatrix} 9 & 10 \\ 11 & 12 \end{bmatrix} \quad , \quad x = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad , \quad y = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \quad , \quad \gamma = 2. \tag{1}$$

Four Properties Of Matrix Multiplication

We begin with the **associativity** property. Namely

$$(AB)C = A(BC). \tag{2}$$

We first begin with $(AB)C$

$$(AB)C = \left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \right) \begin{bmatrix} 9 & 10 \\ 11 & 12 \end{bmatrix} = \left(\begin{bmatrix} 1(5) + 2(7) & 1(6) + 2(8) \\ 3(5) + 4(7) & 3(6) + 4(8) \end{bmatrix} \right) \begin{bmatrix} 9 & 10 \\ 11 & 12 \end{bmatrix}$$

$$= \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix} \begin{bmatrix} 9 & 10 \\ 11 & 12 \end{bmatrix} = \begin{bmatrix} 19(9) + 22(11) & 19(10) + 22(12) \\ 43(9) + 50(11) & 43(10) + 50(12) \end{bmatrix} = \begin{bmatrix} 413 & 454 \\ 937 & 1030 \end{bmatrix}.$$

Next we show that the last line of $(AB)C$ will be equal to $A(BC)$

$$A(BC) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \left(\begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \begin{bmatrix} 9 & 10 \\ 11 & 12 \end{bmatrix} \right) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \left(\begin{bmatrix} 7(9) + 6(11) & 5(10) + 6(12) \\ 7(9) + 8(11) & 7(10) + 8(12) \end{bmatrix} \right)$$

$$= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 111 & 122 \\ 151 & 166 \end{bmatrix} = \begin{bmatrix} 1(111) + 2(151) & 1(122) + 2(166) \\ 3(111) + 4(151) & 3(122) + 4(166) \end{bmatrix} = \begin{bmatrix} 413 & 454 \\ 937 & 1030 \end{bmatrix}.$$

From the above results we have demonstrated the property found in (2), namely the **associativity** property.

We next will demonstrate **associativity with scalar multiplication**, namely

$$\gamma(AB) = (\gamma A)B. \tag{3}$$

Using the values for matrices and scalar values in (1) for $\gamma(AB)$ we have

$$\gamma(AB) = (2) \left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \right) = 2 \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix} = \begin{bmatrix} 38 & 44 \\ 86 & 100 \end{bmatrix}.$$

We now move on to calculating $(\gamma A)B$, for this we have

$$(\gamma A)B = \left((2) \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \right) \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 2(5) + 4(7) & 2(6) + 4(8) \\ 6(5) + 8(7) & 6(6) + 8(8) \end{bmatrix} = \begin{bmatrix} 28 & 44 \\ 86 & 100 \end{bmatrix}.$$

From we the above two results, we have shown the property in (3), namely the **associativity with scalar multiplication**.

We will next show the **distributivity with addition** property. Namely

$$A(B + C) = AB + AC \tag{4}$$

$$(A + B)C = AC + BC. \tag{5}$$

We first begin by showing the property depicted by (4). The LHS of (4) is

$$A(B + C) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \left(\begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} + \begin{bmatrix} 9 & 10 \\ 11 & 12 \end{bmatrix} \right) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \left(\begin{bmatrix} 5+9 & 6+10 \\ 7+11 & 8+12 \end{bmatrix} \right) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 14 & 16 \\ 18 & 20 \end{bmatrix}$$

$$= \begin{bmatrix} 1(14) + 2(18) & 1(16) + 2(20) \\ 3(14) + 4(18) & 3(16) + 4(20) \end{bmatrix} = \begin{bmatrix} 50 & 56 \\ 114 & 128 \end{bmatrix}.$$

The RHS of (4) is

$$AB + BC = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 9 & 10 \\ 11 & 12 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix} + \begin{bmatrix} 31 & 34 \\ 71 & 78 \end{bmatrix} = \begin{bmatrix} 19+31 & 22+34 \\ 43+71 & 50+78 \end{bmatrix}$$

$$= \begin{bmatrix} 50 & 56 \\ 114 & 128 \end{bmatrix}.$$

From the results above we have shown the property found in (4). We now move on to showing (5). The LHS of (5) is

$$(A + B)C = \left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \right) \begin{bmatrix} 9 & 10 \\ 11 & 12 \end{bmatrix} = \left(\begin{bmatrix} 1+5 & 2+6 \\ 3+7 & 4+18 \end{bmatrix} \right) \begin{bmatrix} 9 & 10 \\ 11 & 12 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 10 & 22 \end{bmatrix} \begin{bmatrix} 9 & 10 \\ 11 & 12 \end{bmatrix}$$

$$= \begin{bmatrix} 6(9) + 8(11) & 6(10) + 8(12) \\ 10(9) + 22(11) & 10(10) + 22(12) \end{bmatrix} = \begin{bmatrix} 142 & 156 \\ 222 & 244 \end{bmatrix}.$$

The RHS of (5) is then

$$AC + BC = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 9 & 10 \\ 11 & 12 \end{bmatrix} + \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \begin{bmatrix} 9 & 10 \\ 11 & 12 \end{bmatrix} = \begin{bmatrix} 31 & 34 \\ 71 & 78 \end{bmatrix} + \begin{bmatrix} 111 & 122 \\ 151 & 166 \end{bmatrix} = \begin{bmatrix} 31+111 & 34+122 \\ 71+151 & 78+166 \end{bmatrix}$$

$$= \begin{bmatrix} 142 & 156 \\ 222 & 244 \end{bmatrix}.$$

From the results above we have shown the property found in (5). The results from the previous four calculations have successfully shown the **distributivity with addition**.

The last of the four **properties of matrix multiplication** is the **transpose of product**. Namely

$$(AB)^T = B^T A^T. \tag{6}$$

Using the definitions found in (1) we will now show (6). Using these matrices we now start with the LHS of (6). The LHS is then

$$(AB)^T = \left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \right)^T = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}^T = \begin{bmatrix} 19 & 43 \\ 22 & 50 \end{bmatrix}.$$

The RHS of (6) is then

$$B^T A^T = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}^T \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^T = \begin{bmatrix} 5 & 7 \\ 6 & 8 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 5(1) + 7(2) & 5(3) + 7(4) \\ 6(1) + 8(2) & 6(3) + 8(4) \end{bmatrix} = \begin{bmatrix} 19 & 43 \\ 22 & 50 \end{bmatrix}.$$

From the previous two results we have adequately shown the property depicted in (6). Therefore we have shown the **transpose of product**. We will now move on to other matrix multiplication properties.

Inner Product And Matrix-Vector Products

The **inner product and matrix-vector products** is

$$y^T(Ax) = (y^TA)x = (A^Ty)^Tx. \tag{7}$$

Using the definitions in (1) we can show these properties. These properties with the defined examples are

$$y^T(Ax) = \begin{bmatrix} 3 & 4 \end{bmatrix} \left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right) = \begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} 1(1) + 2(2) \\ 3(1) + 4(2) \end{bmatrix} = \begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 11 \end{bmatrix} = 3(5) + 4(11) = 15 + 44 = 59$$

$$(y^TA)x = \left(\begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \right) \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3(1) + 3(4) & 3(2) + 4(4) \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 15 & 22 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 15 + 44 = 59$$

$$(A^Ty)^Tx = \left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^T \begin{bmatrix} 3 \\ 4 \end{bmatrix} \right)^T \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \left(\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} \right)^T \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1(3) + 3(4) \\ 2(3) + 4(4) \end{bmatrix}^T \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 15 \\ 22 \end{bmatrix}^T \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} 15 & 22 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 15(1) + 22(2) = 15 + 44 = 59.$$

From the above, we can see that all three expressions in (7) have been satisfied. Thus we have adequately shown the **inner product and matrix-vector products**.

Block Matrix Multiplication

To demonstrate the **block matrix multiplication** we must define some more matrices along those found in (1)j. These new matrices are

$$D = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \ , \ E = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \ , \ F = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \ , \ G = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \ , \ H = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \tag{8}$$

The identity for **block matrix multiplication** is

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} E & F \\ G & H \end{bmatrix} = \begin{bmatrix} AE + BG & AF + BH \\ CE + DG & CF + DH \end{bmatrix}. \tag{9}$$

For brevity's sake I am going to abbreviate some of the matrix multiplications for this property. Beginning with the LHS of (9) we have

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} E & F \\ G & H \end{bmatrix} = \begin{bmatrix} 1 & 2 & 5 & 6 \\ 3 & 4 & 7 & 8 \\ 9 & 10 & 1 & 1 \\ 11 & 12 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 5 & 1 & 6 \\ 7 & 7 & 3 & 8 \\ 19 & 1 & 9 & 1 \\ 23 & 0 & 11 & 0 \end{bmatrix}.$$

We will now show the RHS of (9). To do this, we will calculate the multiplications of each sub matrix. These multiplications are then

$$AE + BG = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 7 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 5 \\ 0 & 7 \end{bmatrix} = \begin{bmatrix} 3 & 5 \\ 7 & 7 \end{bmatrix}$$

$$AF + BH = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 6 \\ 0 & 8 \end{bmatrix} = \begin{bmatrix} 1 & 6 \\ 2 & 8 \end{bmatrix}$$

$$CE + DG = \begin{bmatrix} 9 & 10 \\ 11 & 12 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 19 & 0 \\ 23 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 19 & 1 \\ 23 & 0 \end{bmatrix}$$

$$CF + DH = \begin{bmatrix} 9 & 10 \\ 11 & 12 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 9 & 0 \\ 11 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 9 & 1 \\ 11 & 0 \end{bmatrix}.$$

Stitching all of these results together we now have

$$\begin{bmatrix} AE + BG & AF + BH \\ CE + DG & CF + DH \end{bmatrix} = \begin{bmatrix} 3 & 5 & 1 & 6 \\ 7 & 7 & 3 & 8 \\ 19 & 1 & 9 & 1 \\ 23 & 0 & 11 & 0 \end{bmatrix}.$$

The previous results have shown and verified the **block matrix multiplication** property with our given examples.

Column Interpretation Of Matrix-Matrix Product

The **column interpretation of matrix-matrix product** is as follows

$$AB = A \begin{bmatrix} b_1 & b_2 & \dots & b_n \end{bmatrix} = \begin{bmatrix} Ab_1 & Ab_2 & \dots & Ab_n \end{bmatrix}. \tag{10}$$

Using the matrices from (1) we can demonstrate (10) as

$$AB = \begin{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 7 \end{bmatrix} & \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 6 \\ 8 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1(5) + 2(7) \\ 3(5) + 4(7) \end{bmatrix} & \begin{bmatrix} 1(6) + 2(8) \\ 3(6) + 4(8) \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 19 \\ 43 \end{bmatrix} & \begin{bmatrix} 22 \\ 50 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}.$$

In the above, we have correctly demonstrated the **column interpretation of matrix-matrix product** property by multiplying the matrix A by each column of B.

Multiple Sets Of Linear Equations

The **multiple sets of linear equation** is as follows

$$AX = B. \tag{11}$$

The matrix X is just a generalized matrix of variables. Using the matrices in (1) (11) can be expressed as

$$AX = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_3 & x_2 + 2x_4 \\ 3x_1 + 4x_3 & 3x_2 + 4x_4 \end{bmatrix} = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = B.$$

The results from the above result in the following linear equations

$$x_1 + 2x_3 = 5$$
$$x_2 + 2x_4 = 6$$
$$3x_1 + 4x_3 = 7$$
$$3x_2 + 4x_4 = 8.$$

The goal of (11) is to solve for X, using the matrices defined in (1) we can then solve for X with

$$X = A^{-1}B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^{-1} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} -3 & -4 \\ 4 & 5 \end{bmatrix}$$

The above then means the variables for X are

$$x_1 = -3$$
$$x_2 = -4$$
$$x_3 = 7$$
$$x_4 = 5.$$

We have adequately shown the **multiple sets of linear equation** property.

Inner Product Representation

The **inner product representation** is

$$AB = \begin{bmatrix} a_1^T b_1 & a_1^T b_2 & \dots & a_1^T b_n \\ a_2^T b_1 & a_2^T b_2 & \dots & a_2^T b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_m^T b_1 & a_m^T b_2 & \dots & a_m^T b_n \end{bmatrix}. \tag{12}$$

Using the matrices from (1) we can demonstrate (12). This demonstration is then

$$a_1^T b_1 = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 5 \\ 7 \end{bmatrix} = 1(5) + 2(7) = 5 + 14 = 19$$

$$a_1^T b_2 = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ 8 \end{bmatrix} = 1(6) + 2(8) = 6 + 16 = 22$$

$$a_2^T b_1 = \begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 7 \end{bmatrix} = 3(5) + 4(7) = 15 + 28 = 43$$

$$a_2^T b_2 = \begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} 6 \\ 8 \end{bmatrix} = 3(6) + 4(8) = 18 + 32 = 50.$$

Stitching this all together we then have

$$AB = \begin{bmatrix} a_1^T b_1 & a_1^T b_2 \\ a_2^T b_1 & a_2^T b_2 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}.$$

From the previous result we have adequately demonstrated **inner product representation**.

Gram Matrix

The **Gram matrix** is as follows

$$G = A^T A = \begin{bmatrix} a_1^T a_1 & a_1^T a_2 & \ldots & a_1^T a_n \\ a_2^T a_1 & a_2^T a_2 & \ldots & a_2^T a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_m^T a_1 & a_m^T a_2 & \ldots & a_m^T a_n \end{bmatrix}. \tag{13}$$

Using the matrix defined in (1) the **Gram matrix** representation of (13) is then

$$a_1^T a_1 = \begin{bmatrix} 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = 1(1) + 3(3) = 1 + 9 = 10$$

$$a_1^T a_2 = \begin{bmatrix} 1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = 1(2) + 3(4) = 2 + 12 = 14$$

$$a_2^T a_2 = \begin{bmatrix} 2 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = 2(1) + 4(3) = 2 + 12 = 14$$

$$a_2^T b_2 = \begin{bmatrix} 2 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = 2(2) + 4(4) = 4 + 16 = 20.$$

Stitching this all together we then have

$$G = A^T A = \begin{bmatrix} a_1^T a_1 & a_1^T a_2 \\ a_2^T a_1 & a_2^T a_2 \end{bmatrix} = \begin{bmatrix} 10 & 14 \\ 14 & 20 \end{bmatrix}.$$

With the previous results we have adequately shown the **Gram matrix** representation.

Outer Product Representation

The **outer product representation** is as follows

$$AB = a_1 b_1^T + \cdots + a_p b_p^T. \tag{14}$$

Using the matrices defined in (1) the **outer product representation** defined in (14) is

$$a_1 b_1^T = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \begin{bmatrix} 5 & 6 \end{bmatrix} = \begin{bmatrix} 1(5) & 1(6) \\ 3(5) & 3(6) \end{bmatrix} = \begin{bmatrix} 5 & 6 \\ 15 & 18 \end{bmatrix}$$

$$a_2 b_2^T = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \begin{bmatrix} 7 & 8 \end{bmatrix} = \begin{bmatrix} 2(7) & 2(8) \\ 4(7) & 4(8) \end{bmatrix} = \begin{bmatrix} 14 & 16 \\ 28 & 32 \end{bmatrix}.$$

Stitching this all together we then have

$$AB = a_1 b_1^T + a_2 b_2^T = \begin{bmatrix} 5 & 6 \\ 15 & 18 \end{bmatrix} + \begin{bmatrix} 14 & 16 \\ 28 & 32 \end{bmatrix} = \begin{bmatrix} 5 + 14 & 6 + 16 \\ 15 + 28 & 18 + 32 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}.$$

From the above result we have adequately shown the **outer product representation** of a matrix.

Summary

Many of these examples 'recycled' previous calculations to save space.

Problem 1 Summary

Procedure

- For each property, generate an example for each one and show how the property is carried out arithmetically.

Key Concepts

- **Four Properties of Matrix Multiplication:**

    1. **Associativity**: Demonstrated using $(AB)C = A(BC)$ where A, B, and C are matrices.
    2. **Associativity with Scalar Multiplication**: Shown through $\lambda(AB) = (\lambda A)B$.
    3. **Distributivity with Addition**: Verified using $A(B+C) = AB+AC$ and $(A+B)C = AC+BC$.
    4. **Transpose of Product**: Illustrated by $(AB)^T = B^T A^T$.

- **Inner Product and Matrix-Vector Products:**

    - Demonstrated the property $y^T(Ax) = (y^T A)x = (A^T y)^T x$.
    - Used simple matrices to show the equality of these expressions.

- **Block Matrix Multiplication:**

    - Showcased the multiplication of block matrices.
    - Illustrated using block matrices and their multiplication property.

- **Column Interpretation of Matrix-Matrix Product:**

    - Demonstrated that $AB = A[b_1 \ b_2 \ \ldots \ b_n] = [Ab_1 \ Ab_2 \ \ldots \ Ab_n]$.
    - Provided a clear example using specific matrices.

- **Multiple Sets of Linear Equations:**

    - Explored the concept using the matrix equation $AX = B$.
    - Solved for matrix X using given matrices.

- **Inner Product Representation:**

    - Illustrated the inner product representation of matrix multiplication.
    - Used example matrices to demonstrate the property.

- **Gram Matrix:**

    - Explored the concept of the Gram matrix $G = A^T A$.
    - Computed the Gram matrix using provided matrices.

- **Outer Product Representation:**

    - Demonstrated the outer product representation of matrix multiplication.
    - Showed how a matrix product can be expressed as a sum of outer products.

Variations

- We could be asked to show examples of different properties.

    - In this case we would perform the operations for these properties similar to what we did in this problem.

Problem 2

Problem Statement

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).

Original Question:

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?

Solution

For this problem I will be using **method 4**.

Explanation:

In its general form, the Gram matrix is defined as

$$G = A^T A = \begin{bmatrix} a_1^T a_1 & a_1^T a_2 & \dots & a_1^T a_n \\ a_2^T a_1 & a_2^T a_2 & \dots & a_2^T a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_m^T a_1 & a_m^T a_2 & \dots & a_m^T a_n \end{bmatrix}. \tag{1}$$

In QR factorization, we can express a condition for the column vectors of A such that they are orthonormal. Simply put this relationship is

$$A^T A = I. \tag{2}$$

Using the definition of a Gram matrix in (1) with the identity found in (2), we use $A = QR$ in (1). Doing so we have

$$G = A^T A = (QR)^T (QR) = Q^T R^T Q R = R^T Q^T Q R = R^T I R = R^T R. \tag{3}$$

From (3) we can see that the Gram matrix G of A is the same as the Gram matrix of R. In (3) we used the orthonormal condition of Q in (2) to simplify the expression to demonstrate the result.

Because the Gram matrix of A is the same as the Gram matrix of R, this constitutes that the angles between the column vectors in A have to be the same as the angles between the column vectors of R. Namely,

$$\angle(a_i, a_j) = \angle(r_i, r_j). \tag{4}$$

We would not be able to arrive to the conclusion of the Gram matrix of A found in (3) otherwise.

Problem 2 Summary

Procedure

- Use the definition of the Gram matrix.

- Insert (2) into (1) and go through the algebra to get the final result.

Key Concepts

- **Problem Statement:**

  - The problem focuses on understanding the relationship between the Gram matrix of a matrix A with linearly independent columns and the Gram matrix of R in the QR factorization $A = QR$.

  - It also explores the implications of this relationship on the angles between the columns of A and R.

- **Concepts:**

  - **Gram Matrix**: Defined as $G = A^T A$, where A is a matrix and A^T is its transpose.

  - **QR Factorization**: Decomposition of matrix A into a product QR, where Q has orthonormal columns and R is upper triangular.

- **Relationship Between Gram Matrices:**

  - Using $A = QR$, the Gram matrix of A is derived as:

  $$G = (QR)^T(QR) = R^T Q^T QR = R^T R$$

  - This demonstrates that the Gram matrix of A is the same as the Gram matrix of R.

  - The relationship is established by utilizing the orthonormal property of Q, where $Q^T Q = I$.

- **Angles Between Column Vectors:**

  - The equality of the Gram matrices implies that the angles between the column vectors of A are the same as those between the column vectors of R.

  - This is expressed as $\langle a_i, a_j \rangle = \langle r_i, r_j \rangle$ for column vectors a_i, a_j of A and r_i, r_j of R.

- **Linear Algebra Application:**

  - This problem illustrates the application of linear algebra in understanding matrix factorizations and their properties.

  - It highlights how QR factorization affects the properties of matrices and their geometric interpretations.

Variations

- We could be given a different matrix to demonstrate this operation.

  - We would then go through the same procedure as depicted in this problem for the new matrix.

Problem 3

Problem Statement

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).

Original Question:

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

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

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

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

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 $A_1, \dots, A_k$.

Solution

For this problem I will be using **method 4**.

Explanation:

We know in general that QR factorization is calculated with

$$A = QR. \tag{1}$$

In order to calculate the QR factorization for all submatrices we need to use a block matrix representation. The block representation of A, denoted as $\mathbf{A}$ is then

$$\mathbf{A} = \begin{bmatrix} A_i & A_{:,i+1:k} \end{bmatrix}. \tag{2}$$

In (2), the first submatrix A_i is accounting for all the rows and columns of A. The second submatrix $A_{:,i+1:k}$ is taking all of the rows and the second column up to k for the columns. Consequently we also define to the block matrix of Q as $\mathbf{Q}$ with

$$\mathbf{Q} = \begin{bmatrix} Q_i & Q_{:,i+1:k} \end{bmatrix}. \tag{3}$$

Where the same row and column selection applies to (3) as it did to (2). We now have to work on determining the $i \times i$ block matrix $\mathbf{R}$.

In order to get the desired result of $A_i = Q_i R_i$ in our QR factorization, this means at least one entry in $\mathbf{R}$ must be zero. This entry is specifically $\mathbf{R}_{21}$. Consequently, the entry $\mathbf{R}_{11} = R_i$. Following the same pattern as the definition of $\mathbf{A}$ and $\mathbf{Q}$ we can then say $\mathbf{R}$ is

$$\mathbf{R} = \begin{bmatrix} R_i & R_{1:i,i+1:k} \\ 0 & R_{i+1:k,i+1:k} \end{bmatrix} \tag{4}$$

Stitching (2), (3), and (4) together we then have

$$\begin{bmatrix} A_i & A_{:,i+1:k} \end{bmatrix} = \begin{bmatrix} Q_i & Q_{:,i+1:k} \end{bmatrix} \begin{bmatrix} R_i & R_{1:i,i+1:k} \\ 0 & R_{i+1:k,i+1:k} \end{bmatrix} = \begin{bmatrix} Q_i R_i & Q_i R_{1:i,i+1:k} + Q_{:,i+1:k} R_{i+1:k,i+1:k} \end{bmatrix}. \tag{5}$$

We can see immediately from (5) that the first block matrix is the QR factorization of A_i.

Problem 3 Summary

Procedure

- Illustrate what the matrices A and Q are represented as in block matrix form.

- Perform the matrix multiplication to find the final representation of A.

Key Concepts

- **Problem Statement:**
  - The problem involves understanding the QR factorization of the first i columns of a matrix A.
  - The matrix A is an $n \times k$ matrix with QR factorization $A = QR$, where Q has orthonormal columns and R is upper triangular.
  - Submatrices A_i, Q_i, and R_i are defined, and the goal is to show that $A_i = Q_i R_i$ is the QR factorization of A_i.

- **Matrix Notations:**
  - $A_i = A_{1:n,1:i}$ represents the first i columns of A.
  - $Q_i = Q_{1:n,1:i}$ represents the first i columns of Q.
  - $R_i = R_{1:i,1:i}$ is the $i \times i$ submatrix of R containing its first i rows and columns.

- **Demonstration of QR Factorization:**
  - The demonstration involves using block matrix representation to show that A_i can be expressed as the product of Q_i and R_i.
  - The block representation of A, Q, and R are used to partition these matrices and demonstrate the QR factorization for the submatrix A_i.

- **Solution and Explanation:**
  - The solution is provided through detailed matrix block manipulation and algebraic expressions.
  - The solution concludes that $A_i = Q_i R_i$ is indeed the QR factorization of A_i, showing that the QR factorization of A simultaneously provides the QR factorization of all its leading submatrices.

- **Linear Algebra Concepts:**
  - This problem provides insights into QR factorization and how it applies to submatrices.
  - It emphasizes the use of block matrices in simplifying complex algebraic expressions and understanding matrix factorizations.

Variations

- We could be asked to demonstrate this multiplication with a different form.
  - We would then go through the same procedure to depict the desired form that is requested from us.

Problem 4

Problem Statement

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)

Original Question:

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.

Solution

For this problem I will be using **method 4**.

Explanation:

From page 102 of VMLS, we are told that the order of flops for matrix multiplication is on the order of n^3. In the problem statement we were told that the time to multiply two 1500×1500 matrices took around 0.2 seconds. We are trying to determine the time taken to multiply two matrices together that are twice in size of the original matrix.

If we are trying to estimate the time taken to multiply two matrices that are twice the size of the original matrices (since $3000/1500 = 2$) then this means $n = 2$ in our expression. Taking this into account we can then say

$$n^3 = 2^3 = 8. \tag{1}$$

The result from (1) means that it would take approximately 8 times longer to multiply the 3000×3000 matrices. Since the original time is 0.2 seconds, this means the estimated time taken to multiply the 3000×3000 matrices is roughly

$$8 \cdot 0.2 \text{ seconds} = 1.6 \text{ seconds}. \tag{2}$$

So a rough estimate would be 1.6 seconds.

Problem 4 Summary

Procedure

- Use the result from the smaller matrix to calculate the time taken for the larger matrix with $n = 2$ in n^3.

Key Concepts

- **Problem Statement:**
  - The problem assesses the computational complexity of matrix multiplication in relation to matrix size.
  - Given data: a computer takes approximately 0.2 seconds to multiply two 1500×1500 matrices.
  - The task is to estimate the time it would take for the same computer to multiply two 3000×3000 matrices.

- **Solution Approach:**
  - The solution employs the cubic complexity of matrix multiplication, denoted as n^3, where n represents the matrix size.
  - Since the size of the matrices increases from 1500 to 3000, which is double the size, the complexity factor n is set to 2.

- **Computational Complexity Estimation:**
  - It is deduced that the time taken for matrix multiplication increases by a factor of n^3, hence $2^3 = 8$ in this case.
  - This implies that multiplying 3000×3000 matrices will take approximately 8 times longer than 1500×1500 matrices.

- **Time Calculation:**
  - The estimated time is calculated as 8×0.2 seconds, resulting in 1.6 seconds.
  - The final answer for the estimated time is 1.6 seconds.

- **Linear Algebra and Computer Science Application:**
  - This problem illustrates the application of linear algebra concepts in the field of computer science, specifically in understanding the computational efficiency of matrix operations.
  - It highlights the significance of matrix dimensions in determining the computational cost of matrix multiplication.

Variations

- We could be given a different dimension for the new matrix.
  - We would then figure out the n value for this new matrix and feed it into the formula to find the time.