8 Matrix Multiplication

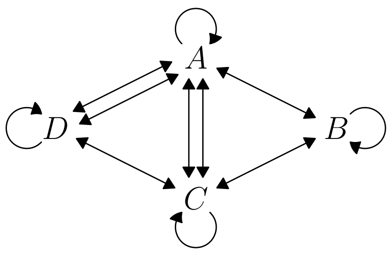


Figure 1: Four town transportation scenario.

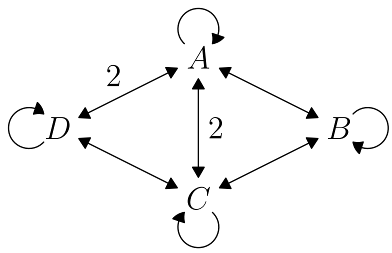


Figure 2: The scenario, with numbers instead of duplicate lines.

You've all seen a bunch of numbers organized in a table. Sometimes a table is just a table, but sometimes we will call it a **matrix**.

What makes a matrix different from a table? One important difference is that we can meaningfully *multiply* matrices. The purpose of this lesson is to explain why the matrix multiplication rule makes sense and when it is useful.

Consider a region with four towns named A, B, C, D. There are modes of transport between these towns; a path can go from a town to any town, including the same town. These paths are shown in Figure 1. At each "step," you can take any path from one town to the next. For example, you might start on A, then take either of the two paths to D. But you cannot start on D and go directly to B. When there is more than one path between towns, we could also just draw a line and label it with a number: this is shown in Figure 2. Note how each town has a "path" going to itself. Taking this means you don't go anywhere.

Let us consider a **transportation matrix** in this scenario. The number in the i^{th} row and j^{th} column of the matrix A, which we'll call a_{ij}, gives the number of ways to walk directly from town i to town j. For example, $a_{42} = a_{24} = 0$, because there are no ways to get from B to D or D to B.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} = \begin{array}{cc} & \text{to} \\ & \begin{array}{cccc} A & B & C & D \end{array} \\ \text{from}\begin{array}{c} A \\ B \\ C \\ D \end{array} & \begin{bmatrix} 1 & 1 & 2 & 2 \\ 1 & 1 & 1 & 0 \\ 2 & 1 & 1 & 1 \\ 2 & 0 & 1 & 1 \end{bmatrix} \end{array}$$

This matrix is **symmetric**, meaning there are no one-way paths. In mathematical terms, we have $a_{ij} = a_{ji}$ for all valid i, j. In visual terms, the matrix is symmetric about the **main diagonal**, going from top-left to bottom-right. Furthermore, the main diagonal is all 1s, because we allow staying in the town you start in.

Suppose there's a shuttle bus that only goes one way, from town A to B to C to D and then back to A again. The transportation matrix for this scenario—again, allowing staying still—is shown in Figure 3.

$$B = \begin{array}{cc} & \text{to} \\ & \begin{array}{cccc} A & B & C & D \end{array} \\ \text{from}\begin{array}{c} A \\ B \\ C \\ D \end{array} & \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

Figure 3: Transportation matrix B.

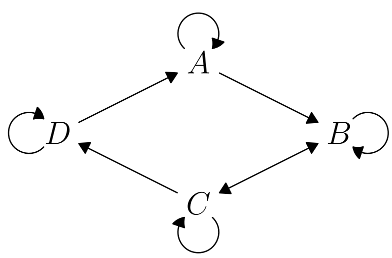


Figure 4: Graph of matrix B.

Because there are one-way connections, B is not symmetric. For example, $b_{12} = 1$, but $b_{21} = 0$. The graph for this matrix is therefore "directed"; it has arrows indicating the direction of each path. This is shown in Figure 4.

Now, suppose you wanted to know the total number of ways to go from town to town in one step, by path or by bus. To find the total, you add the matrices in the obvious way: term by term, or $(a + b)_{ij} = a_{ij} + b_{ij}$. We can go directly $C \to D$, but not directly $D \to B$. This is shown in Figure 5.

$$A{+}B = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 1 & 1 & 1 & 0 \\ 2 & 1 & 1 & 1 \\ 2 & 0 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 2 & 2 \\ 1 & 2 & 2 & 0 \\ 2 & 1 & 2 & 2 \\ 3 & 0 & 1 & 2 \end{bmatrix}$$

Figure 5: Matrix addition of A and B.

$$A + B = \begin{bmatrix} 1 & 2 & 2 & 2 \\ 1 & 1 & 2 & 0 \\ 2 & 1 & 1 & 2 \\ 3 & 0 & 1 & 1 \end{bmatrix}$$

Figure 6: $A + B$, with 1s on the diagonal.

Okay, so that's a little silly: we've counted two different ways to stay still, namely "not taking a path" and "not going anywhere on the bus." We should rewrite the matrix, putting ones on the diagonal, as in Figure 6. Despite this minor issue, it's still true in general that this most naïve way of adding matrices is also the most convenient and useful. Just don't blindly use an operation without considering its meaning!

But now comes a surprise: the most useful way to multiply matrices is not the obvious way. Why not? You'll see several different examples in the coming weeks. For now, think: what would it mean in terms of transportation if we just multiplied corresponding numbers like $a_{13}b_{13}$?[7] It would be meaningless, as far as I can tell.

Instead, we want multiplication of the two matrices B and A to represent taking one step by walking and then one step by bus. Similarly, multiplication of matrix A by itself will represent the number of ways to go from town to town in two steps by walking.

So, what rule of matrix multiplication will make that happen? To get from town A to C in two steps, for example, we have to go from town A to one of the four towns, then from that town to town C. The total number of ways to do this is

$$\underbrace{a_{11}}_{A\to A} \cdot \underbrace{a_{13}}_{A\to C} + \underbrace{a_{12}}_{A\to B} \cdot \underbrace{a_{23}}_{B\to C} + \underbrace{a_{13}}_{A\to C} \cdot \underbrace{a_{33}}_{C\to C} + \underbrace{a_{14}}_{A\to D} \cdot \underbrace{a_{43}}_{D\to C} = \sum_{j=1}^{4} a_{1j}a_{j3}.$$

And that's how we'll eventually define matrix multiplication. We can say that to determine the ij entry of the product XY of matrices X and Y, use the following formula:

$$(XY)_{ij} = x_{i1}y_{1j} + x_{i2}y_{2j} + \cdots + x_{in}y_{nj} = \sum_{k=1}^{n} x_{ik}y_{kj},$$

where X has n columns and Y has n rows. This definition is fine and dandy if you're, say, programming a computer to do matrix multiplication, but we should find a more intuitive way to interpret it.

Suppose we're multiplying two matrices X and Y. For convenience, let's make them 2×3 [8] and 3×2:

$$\underbrace{\begin{bmatrix} 1 & 2 & 3 \\ \boxed{4 \quad 5 \quad 6} \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} \boxed{\begin{matrix} 7 \\ 9 \\ 11 \end{matrix}} & \begin{matrix} 8 \\ 10 \\ 12 \end{matrix} \end{bmatrix}}_{Y} = \begin{bmatrix} 1\cdot 7 + 2\cdot 9 + 3\cdot 11 & 1\cdot 8 + 2\cdot 10 + 3\cdot 12 \\ 4\cdot 7 + 5\cdot 9 + 6\cdot 11 & 4\cdot 7 + 9\cdot 10 + 11\cdot 12 \end{bmatrix} = \begin{bmatrix} 58 & 64 \\ \boxed{139} & 250 \end{bmatrix}.$$

Observe the boxed numbers. To get 139, we multiplied the boxed row and column term-by-term. That is, we did $4 \cdot 7 + 5 \cdot 9 + 6 \cdot 11$. We can think of this as the dot product of two vectors:

$$\langle 4, 5, 6\rangle \cdot \langle 7, 9, 11\rangle = 139.$$

In our example, to find the top left number a_{11}, we'd do $\langle 1, 2, 3\rangle \cdot \langle 7, 9, 11\rangle$. In general, to find $(XY)_{ij}$, we find the dot product of the ith **row vector** of X and jth **column vector** of Y. More intuitively, to compute the entry in the ith row and jth column of XY, run your left hand across the ith row of matrix X and your right hand down the jth column of matrix Y, multiplying the elements and adding the products as you go along.

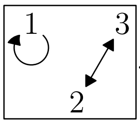
With the ability to multiply matrices more easily, let's try some problems.

1. The 3-post snap group can be represented by a set of graphs, each with three towns. The posts are the towns and the elastic bands are the roads. An example is shown.

[7] This "obvious" product is known as the Hadamard product.

[8] Note that the first dimension is rows and the second is columns, as is the usual order.

$$
A = \text{from}\ \begin{array}{c c} & \begin{array}{ccc} & \text{to} & \\ 1 & 2 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{array}\right] \end{array} \longleftrightarrow
$$



.

(a) Draw the graphs and transportation matrices for this group.

(b) Try a few multiplications and notice the isomorphism to the snap group.

2. Using 3×3 matrices A and B from this section, compute:

   (a) $AA = A^2$ (b) AB (c) BA (d) B^2

   (e) Which one (AB or BA) represents taking a step by walking, then by bus?

   (f) Use your calculator to check your computations of A^2, AB, BA, and B^2.

3. Write a 3×3 matrix T that shows the following scenario: You can go from town B to C, C to D, and D to B by train, in exactly one way each, and not backwards.

   (a) Why can't you add this matrix to matrices A or B?

   (b) Rewrite matrix T so that it *can* be meaningfully added to matrices A and B. What did you do to its dimensions?

It's time for a review of **sigma notation**! Sigma notation represents a sum. It is defined as

$$
S = \sum_{k=m}^{n} f(k) = f(m) + f(m+1) + \cdots + f(n),
$$

for function f and integers m, n. k is the index over which the summation is taking place. It takes on all integer values between m and n, inclusive. You might read the definition as "The summation of f of k from k equals m to n equals S."

4. Evaluate the following sums:

   (a) $\sum_{k=1}^{4} k$ (b) $\sum_{k=0}^{5} k^2$ (c) $\sum_{k=1}^{10} 3$ (d) $\sum_{k=1}^{n} k$ (e) $\sum_{k=1}^{n} n$ (f) $\sum_{k=1}^{n} 1$

5. The matrix C^T whose rows are the same as the respective columns of matrix C is called the **transpose** of C. For example,

$$
C = \left[\begin{array}{cc} 1 & 2 \\ 3 & 4 \end{array}\right], \; C^T = \left[\begin{array}{cc} 1 & 3 \\ 2 & 4 \end{array}\right].
$$

   (a) Let the elements of C be c_{ij} and the elements of C^T be c'_{ij}. Write a formula for C^T in terms of these elements. That is, $c'_{ij} =$ ___?

   (b) Write $\left[\begin{array}{ccc} 2 & 1 & 5 \\ 4 & -2 & 0 \end{array}\right]^T$.

6. Fill in the blanks: Multiplying an $m \times n$ matrix by a(n) ___ $\times\, k$ matrix gives a(n) ___ $\times$ ___ matrix.

7. Dogs can eat cats, rats, or mice; cats can eat rats or mice; rats can eat mice.

   (a) Make a matrix E showing what can eat what.

   (b) Draw a directed graph.

   (c) Calculate and interpret E^2, E^3, E^4.

The following table shows the amount of each ingredient a bakery uses in making one batch of sourdough bread and biscuits. Of course, the units vary depending on the ingredient. Let's call this matrix S for sourdough.

$$
\begin{array}{l|cccccccc}
\text{Sourdough} & \text{Flour} & \text{Starter} & \text{Yeast} & \text{Water} & \text{Salt} & \text{Soda} & \text{Sugar} & \text{Butter} \\
\hline
\begin{array}{l}\text{Bread}\\ \text{Biscuits}\end{array} & \multicolumn{8}{c}{\left[\begin{array}{cccccccc} 5 & 1 & 0 & \frac{4}{3} & 1 & 1 & 0 & 0 \\ 5 & 1 & 1 & \frac{5}{4} & \frac{3}{4} & 0 & \frac{1}{3} & 2 \end{array}\right]}
\end{array} = S
$$

The bakery wants to know how much the ingredients cost for one batch of bread, and how much for one batch of biscuits. The unit cost of the ingredients is given in the following table. Let's call this matrix C for cost.

$$
\begin{array}{l|cccccccc}
 & \text{Flour} & \text{Starter} & \text{Yeast} & \text{Water} & \text{Salt} & \text{Soda} & \text{Sugar} & \text{Butter} \\
\hline
\text{\$ per unit} & \multicolumn{8}{c}{\left[\begin{array}{cccccccc} 5 & 20 & 10 & 0 & 1 & 2 & 5 & 12 \end{array}\right]}
\end{array} = C
$$

8. (a) Unfortunately, if you try to multiply S and C as given, it won't work. Why not?
   (b) What do you need to do to C so they can be multiplied? Explain the dimensions of each matrix.
   (c) Once you've fixed matrix C, do the multiplication. What are the dimensions of your answer?
9. Matrix multiplication is not necessarily commutative, even when the dimensions of the matrices suggest it might be. How do we know? Be specific.
10. Matrix multiplication is associative, though. Prove that $(PX)T = P(XT)$ for

$$P = \begin{bmatrix} m & n \\ p & q \end{bmatrix}, \; X = \begin{bmatrix} x & y \\ z & w \end{bmatrix}, \; Y = \begin{bmatrix} r & s \\ t & u \end{bmatrix}.$$

11. Prove that matrix multiplication is distributive: $P(X + T) = PX + PT$.
12. Under what conditions does $PX = XP$? Don't worry if you get some messy equations in your answer.
13. Cook's Seafood Restaurant in Menlo Park sells fish and chips. The Captain's order is two pieces of fish and one order of chips, while the Regular order is one piece of fish and one order of chips.
   (a) Write a matrix representing these facts, with clear labels on your rows and columns.
   (b) The restaurant management estimates their cost at 0.75 for each piece of fish and 0.50 for each order of chips. Represent this as a matrix, then use matrix multiplication to calculate the cost of the two possible orders.
   (c) For a party, Cook's provides 10 Captain's orders and 5 Regular orders. Write this as a matrix and use matrix multiplication to find how many pieces of fish and orders of chips are provided.
   (d) Now use matrix multiplication to find out the cost of the party.
14. We will find coefficient matrices to be particularly useful for solving systems of linear equations. For instance,

$$\begin{cases} 3x + 4y &= 5 \\ 6x + 4y &= 8 \end{cases} \quad \longleftrightarrow \quad \begin{bmatrix} 3 & 4 \\ 6 & 7 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 8 \end{bmatrix}.$$

Rewrite

$$\begin{cases} 2x + 3y + 4z &= 5 \\ 5x - 4y + 2z &= 2 \\ x + 2y &= 7 \end{cases}$$

as a matrix equation in this way.

15. (a) What is the transpose of the 3×3 matrix M from the previous problem?
   (b) Use M^T to rewrite the system in the previous problem.
   (c) What is the transpose of the transpose matrix, $(M^T)^T$?