5.1 Matrix-Matrix Multiplication

Matrix-matrix multiplication is another powerful tool for modeling linear systems, which we will discuss further as the class progresses. As an example, two matrices A and B in $\mathbb{R}^{2\times 2}$ can be multiplied as follows:

$$\underset{\text{A}}{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}} \underset{\text{B}}{\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}} = \underset{\text{AB}}{\begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}}$$

Computationally, matrix-matrix multiplication involves multiplying each row vector in A with each column vector in B, starting from the top row of matrix A and leftmost column of matrix B. Effectively, the left matrix is multiplied by each column vector in the second matrix to produce a new column of AB. Why columns and not rows? That's just convention. But this does lead to an important point about the dimensions of matrix-matrix multiplication.

To left-multiply a matrix B by another matrix A, the number of columns in A must equal the number of rows in B. Otherwise, the product $A \times B$ cannot be calculated. Moreover, if A is an $m \times n$ matrix and B is $n \times p$, the product $A \times B$ will have dimensions $m \times p$. A visual illustration of this can be seen here, where the left matrix is broken up into m row vectors and the right matrix is represented as p column vectors:

$$\begin{bmatrix} ---- \vec{r}_1^T ---- \\ ---- \vec{r}_2^T ---- \\ \vdots \\ ---- \vec{r}_m^T ---- \end{bmatrix} \begin{bmatrix} | & | & & | \\ \vec{c}_1 & \vec{c}_2 & \cdots & \vec{c}_p \\ | & | & & | \end{bmatrix} = \begin{bmatrix} \vec{r}_1^T\vec{c}_1 & \vec{r}_1^T\vec{c}_2 & \cdots & \vec{r}_1^T\vec{c}_p \\ \vec{r}_2^T\vec{c}_1 & \vec{r}_2^T\vec{c}_2 & \cdots & \vec{r}_2^T\vec{c}_p \\ \vdots & \vdots & & \vdots \\ \vec{r}_m^T\vec{c}_1 & \vec{r}_m^T\vec{c}_2 & \cdots & \vec{r}_m^T\vec{c}_p \end{bmatrix}$$

In order for the inner product $\vec{r}_i^T\vec{c}_j$ to be defined, each row vector ($\vec{r}_i^T$) must have the same number of entries as each column vector ($\vec{c}_j$). As a result, matrix-matrix multiplication is typically not commutative — $A \times B$ does not necessarily equal $B \times A$. In fact, both quantities can only be calculated if the number of rows in A equals the number of columns in B **and** the number of rows in B equals the number of columns in A.

To illustrate this, consider the following example of taking the product of two 2×2 matrices.

Example 5.1 (Matrix Multiplication):

$$\begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} (2)(1)+(4)(3) & (2)(2)+(4)(4) \\ (3)(1)+(1)(3) & (3)(2)+(1)(4) \end{bmatrix} = \begin{bmatrix} 14 & 20 \\ 6 & 10 \end{bmatrix}$$

Example 5.2 (Matrix Multiplication is Not Commutative!): Above, we mentioned that matrix multiplication does not commute - that is to say, there exist matrices A and B such that $AB \neq BA$. Let's see if we can come up with such an example to verify that assertion.

A natural approach would be to take the matrices from the above example, multiply them in the other order, and see if we get the same answer. Let's try it out! Swapping the order of the matrices, we obtain

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} (1)(2)+(2)(3) & (1)(4)+(2)(1) \\ (3)(2)+(4)(3) & (3)(4)+(4)(1) \end{bmatrix} = \begin{bmatrix} 8 & 6 \\ 18 & 16 \end{bmatrix}$$

As expected, we did not end up with the same result as we did before. Having produced a counterexample, we have therefore proven that matrix multiplication is not generally commutative.

Be aware, however, that there still might (and indeed do!) exist pairs of matrices whose product *is* commutative. All we have shown here is that *not all* pairs of matrices produce the same product when multiplied in the opposite order.

Example 5.3 (Outer Product): For vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$, we have already defined their inner product as $\vec{x}^T \vec{y} = \vec{y}^T \vec{x}$, which is a scalar quantity. What happens if we instead multiply a column-vector by a row-vector? This is just a special case of matrix-matrix multiplication, but it is sometimes called the **outer product** of two vectors.

$$\vec{x}\,\vec{y}^T = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix} = \begin{bmatrix} x_1y_1 & x_1y_2 & \cdots & x_1y_n \\ x_2y_1 & x_2y_2 & \cdots & x_2y_n \\ \vdots & \vdots & & \vdots \\ x_ny_1 & x_ny_2 & \cdots & x_ny_n \end{bmatrix}.$$

We remark that, unlike inner product, the order of $\vec{x}$ and $\vec{y}$ matters in taking an outer product (due to non-commutativity of matrix multiplication). Moreover, it is possible to multiply column vectors times row vectors of different dimensions (another special case of matrix-matrix multiplication), but this is typically not referred to as an outer product.

Example 5.4 (Matrix Multiplication is Associative!): Having seen above that matrix multiplication is not commutative, we might start asking questions about associativity, as well. In particular, is it true that given three matrices A, B, and C, that $(AB)C = A(BC)$? Put differently, does the *grouping* of matrices in a product not matter, if the order is kept the same throughout?

As it turns out, this is true. Unfortunately, a general proof of associativity is tedious and relies on just repeatedly applying the component-wise definition of matrix multiplication. To gain some intuition about associativity, it's better to simply consider an example, such as the following:

$$\begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 7 & 8 \\ 9 & 10 \end{bmatrix} \begin{bmatrix} 11 & 12 \\ 13 & 14 \end{bmatrix}.$$

The above product can be evaluated in two different ways - we will do both, and verify that we get the same answer either way.

Let's first multiply the first two matrices together, before multiplying their product with the third:

$$\begin{aligned}
\left(\begin{bmatrix}3 & 4\\5 & 6\end{bmatrix}\begin{bmatrix}7 & 8\\9 & 10\end{bmatrix}\right)\begin{bmatrix}11 & 12\\13 & 14\end{bmatrix} &= \begin{bmatrix}(3)(7)+(4)(9) & (3)(8)+(4)(10)\\(5)(7)+(6)(9) & (5)(8)+(6)(10)\end{bmatrix}\begin{bmatrix}11 & 12\\13 & 14\end{bmatrix}\\
&= \begin{bmatrix}57 & 64\\89 & 100\end{bmatrix}\begin{bmatrix}11 & 12\\13 & 14\end{bmatrix}\\
&= \begin{bmatrix}(57)(11)+(64)(13) & (57)(12)+(64)(14)\\(89)(11)+(100)(13) & (89)(12)+(100)(14)\end{bmatrix}\\
&= \begin{bmatrix}1459 & 1580\\2279 & 2468\end{bmatrix}.
\end{aligned}$$

Then, let's try multiplying the last two matrices together first, before multiplying the first matrix with that product:

$$\begin{aligned}
\begin{bmatrix}3 & 4\\5 & 6\end{bmatrix}\left(\begin{bmatrix}7 & 8\\9 & 10\end{bmatrix}\begin{bmatrix}11 & 12\\13 & 14\end{bmatrix}\right) &= \begin{bmatrix}3 & 4\\5 & 6\end{bmatrix}\begin{bmatrix}(7)(11)+(8)(13) & (7)(12)+(8)(14)\\(9)(11)+(10)(13) & (9)(12)+(10)(14)\end{bmatrix}\\
&= \begin{bmatrix}3 & 4\\5 & 6\end{bmatrix}\begin{bmatrix}181 & 196\\229 & 248\end{bmatrix}\\
&= \begin{bmatrix}(3)(181)+(4)(229) & (3)(196)+(4)(248)\\(5)(181)+(6)(229) & (5)(196)+(6)(248)\end{bmatrix}\\
&= \begin{bmatrix}1459 & 1580\\2279 & 2468\end{bmatrix},
\end{aligned}$$

which is the same as what we got before!

The fact that three fairly arbitrary matrices exhibit associativity when being multiplied should be a strong hint that matrix multiplication is probably associative - however, it is important to understand that *this is not a proof* of the associativity of matrix multiplication. To prove that matrix multiplication is associative, we'd have to show that *any* triplet of matrices can be multiplied in either order without changing the final answer - showing that it seems to work for particular examples is not sufficient.

Example 5.5 (Water Reservoir): Recall the water reservoir, where applying the matrix to the current distribution of water gives us the next day's distribution:

$$\begin{bmatrix}\frac{1}{2} & \frac{1}{4} & \frac{1}{3}\\ \frac{1}{2} & 0 & \frac{1}{3}\\ 0 & \frac{3}{4} & \frac{1}{3}\end{bmatrix}\begin{bmatrix}x_A\\x_B\\x_C\end{bmatrix} = \begin{bmatrix}x'_A\\x'_B\\x'_C\end{bmatrix}$$

Here, we have three variables, one for each reservoir. We want a function that takes in a water distribution and gives us the water distribution one day later, which is represented by the matrix A.

What if we want the water distribution two days later? We could apply A twice, giving us $A(A\vec{x})$. Alternatively, a key property of matrix multiplication is **associativity**, or $(AB)C = A(BC)$, so we know that $A(A\vec{x}) = (AA)\vec{x}$. Therefore, we can use matrix-matrix multiplication to produce AA:

$$\begin{bmatrix}\frac{1}{2} & \frac{1}{4} & \frac{1}{3}\\ \frac{1}{2} & 0 & \frac{1}{3}\\ 0 & \frac{3}{4} & \frac{1}{3}\end{bmatrix}\begin{bmatrix}\frac{1}{2} & \frac{1}{4} & \frac{1}{3}\\ \frac{1}{2} & 0 & \frac{1}{3}\\ 0 & \frac{3}{4} & \frac{1}{3}\end{bmatrix} = \begin{bmatrix}\frac{3}{8} & \frac{3}{8} & \frac{13}{36}\\ \frac{1}{4} & \frac{3}{8} & \frac{5}{18}\\ \frac{3}{8} & \frac{1}{4} & \frac{13}{36}\end{bmatrix}$$

This operation gives us a single matrix representing two days of water flow. In other words, matrix multiplication implements function composition, and AA represents applying the function A twice.

In algebra, we learned how to manipulate functions of one variable. Linear algebra teaches us how to manipulate linear functions of multiple variables.

> **Additional Resources** For more on matrix-matrix multiplication, read *Strang* pages 61-62, and try Problem Set 2.3.
>
> In *Schuam's*, read pages 30-33 and try Problems 2.4 to 2.11, 2.39 to 2.40, 2.42, 2.44 - 2.49, 2.12 to 2.16, 2.41, 2.43, and 2.72. *Extra: Understand Polynomials in Matrices.*

5.2 Linear Transformations

In an earlier note, we looked at real-valued linear functions on $\mathbb{R}^n$. In particular, we saw that if $f : \mathbb{R}^n \to \mathbb{R}$ was a linear function, then it must be of the form $f(\vec{x}) = \vec{a}^T \vec{x}$ for some vector of coefficients $\vec{a} \in \mathbb{R}^n$.

What if we consider vector-valued linear functions $f : \mathbb{R}^n \to \mathbb{R}^m$? Such a function is often referred to as a **linear transformation**. In order to answer this question, we define precisely the notion of linear transformation.

Definition 5.1 (Linear Transformation): A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is said to be a **linear transformation** (from $\mathbb{R}^n$ to $\mathbb{R}^m$) if it satisfies the following two properties:

- **Homogeneity**. For any $\alpha \in \mathbb{R}$ and $\vec{x} \in \mathbb{R}^n$,

$$f(\alpha\vec{x}) = \alpha f(\vec{x}).$$

- **Additivity** (also known as **Superposition**). For any two vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$,

$$f(\vec{x}+\vec{y}) = f(\vec{x}) + f(\vec{y}).$$

Note that, together, the properties of additivity and superposition are equivalent to f satisfying

$$f(\alpha\vec{x}+\beta\vec{y}) = \alpha f(\vec{x}) + \beta f(\vec{y}) \quad \text{for all } \alpha, \beta \in \mathbb{R} \text{ and } \vec{x}, \vec{y} \in \mathbb{R}^n,$$

which should look familiar.

5.2.1 Matrices as Linear Transformations, and vice versa

So far, we've defined the requirements that a linear transformation must satisfy. But what *is* a linear transformation, really? As it turns out, multiplying a matrix with a column vector is a linear transformation - specifically, the function

$$f_A(\vec{x}) = A\vec{x}$$

is a linear transformation for any matrix A. Typically, we simplify this statement by stating that the matrix A *itself* is a linear transformation, with the matrix used to represent the transformation f_A.

But why is this true? To check if a function is a linear transformation, we simply need to verify that it satisfies the requirements of homogeneity and additivity. Observe that, by the rules of matrix-vector multiplication,

$$f_A(\vec{x}+\vec{y}) = A(\vec{x}+\vec{y}) = A\vec{x}+A\vec{y} = f_A(\vec{x})+f_A(\vec{y})$$
$$f_A(\alpha\vec{x}) = A(\alpha\vec{x}) = \alpha(A\vec{x}) = \alpha f_A(\vec{x}),$$

where $\vec{x}$ and $\vec{y}$ are arbitrary vectors, A is a matrix with the appropriate dimensions, and α is an arbitrary real scalar, so both additivity and homogeneity are satisfied by matrix multiplication. Thus, matrix multiplication is a linear transformation, as we claimed earlier.

On the other hand, if we are given a linear transformation $f : \mathbb{R}^n \to \mathbb{R}^m$, then we can write it as a vector of real-valued linear functions. I.e., $f(\vec{x}) = [f_1(\vec{x}), f_2(\vec{x}), \ldots, f_m(\vec{x})]^T$ for linear functions $f_i : \mathbb{R}^n \to \mathbb{R}$. However, as mentioned previously, every linear function can be represented as an inner product between a vector of coefficients and its input, so we can write $f_i(\vec{x}) = \vec{a}_i{}^T\vec{x}$, $i = 1, \ldots, m$ for an appropriate choice of (coefficient) vectors $\vec{a}_1, \ldots, \vec{a}_m$ in $\mathbb{R}^n$.

In other words, we can express $f(\vec{x}) = A\vec{x}$, where $A \in \mathbb{R}^{m\times n}$ is a matrix with columns $\vec{a}_1, \ldots, \vec{a}_m$.

Thus, we conclude that $m \times n$ matrices explicitly represent the set of linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$. In particular, matrix-vector and matrix-matrix multiplication allow us to explicitly represent linear transformations, and compositions thereof.

One final piece of jargon remains to be introduced - when a linear transformation yields vectors of the same dimension as its input (i.e. if $f(\vec{x})$ has the same dimension as $\vec{x}$) then it is sometimes called a **linear operator**.

5.3 Interpretation: Water Reservoirs and Pumps

Let us revisit our example of matrix-vector multiplication in the context of water reservoirs and water pumps. We are presenting this example because it is vital, as an engineer, to understand the ideas that we are talking about in an intuitive way; this intuition often comes from having a series of examples that make sense. After all, we define mathematical operations the way that we do because these definitions are useful; they don't come out of nowhere. The act of doing mathematics and engineering is often about making definitions and seeing where they lead us, while checking the consistency of these definitions with what we are trying to model in the real world.

For these examples, we will have three water reservoirs, A, B, C. Let's say the initial amounts of water they respectively hold are A_0, B_0, C_0. Next, say we have a system of pumps connecting the reservoirs that move certain amounts of water between the reservoirs every day. We can represent the reservoirs as the following vector, where each element describes how much water is currently in that reservoir:

$$\begin{bmatrix} A \\ B \\ C \end{bmatrix}$$

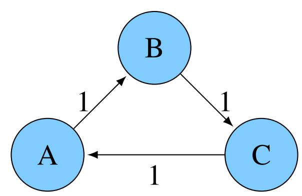
Then, we can represent the system of pumps as a matrix:

$$\begin{bmatrix} P_{A\to A} & P_{B\to A} & P_{C\to A} \\ P_{A\to B} & P_{B\to B} & P_{C\to B} \\ P_{A\to C} & P_{B\to C} & P_{C\to C} \end{bmatrix}$$

Each element $P_{i\to j}$ represents the fraction of water in reservoir i that goes into reservoir j the next day. The matrix acts on the vector just as the pumps act on the reservoirs, performing a transformation — multiplying a vector representing the distribution of water in one day by the pump matrix will give a vector with the distribution of water the next day. We call this matrix a **state transition matrix**. This example can also extend to matrix-matrix multiplication. Both this pumps and reservoirs example and a similar example (PageRank — how search engines can use link information to figure out which pages are important) will show further applications of linear algebra.

5.3.1 Basic Pump

The most basic pump system will move all water from one reservoir into another. Pictorially, we can show this as follows (blue circles are the reservoirs and arrows represent how the pumps move the water):



The corresponding matrix-vector multiplication is:

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix}$$

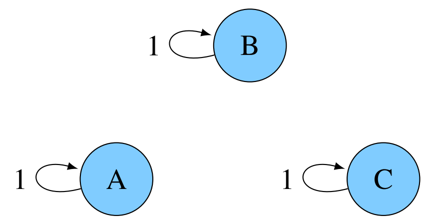
Each time the pumps act on the reservoirs, all of the water in reservoir A flows into reservoir B. All of the water in reservoir B flows into reservoir C. All of the water in reservoir C flows into reservoir A. If A, B, and C all start with the same amount of water, then the pumps acting on the reservoirs would not change the amount of water in each reservoir. As an example, let the amount of water in each reservoir initially be A_0, B_0, C_0. Then we can calculate the amount of water in each reservoir after activating the pumps once (A_1, B_1, C_1) as follows:

$$\begin{bmatrix} A_1 \\ B_1 \\ C_1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} A_0 \\ B_0 \\ C_0 \end{bmatrix} = \begin{bmatrix} C_0 \\ A_0 \\ B_0 \end{bmatrix}$$

5.3.2 Identity Matrix Pump

What happens when your pump system can be represented as the identity matrix? What does that mean? If the initial amounts of water in the reservoirs are represented by the vector $\begin{bmatrix} A_0 \\ B_0 \\ C_0 \end{bmatrix}$, and the identity matrix represents how the pumps move the water, after one activation of the pumps, nothing changes!

$$\begin{bmatrix} A_1 \\ B_1 \\ C_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} A_0 \\ B_0 \\ C_0 \end{bmatrix} = \begin{bmatrix} A_0 \\ B_0 \\ C_0 \end{bmatrix}$$



5.3.3 Drain

Another special matrix is the zero matrix:

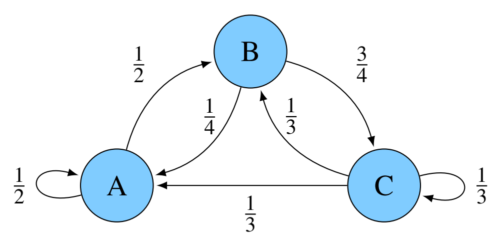
$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

In terms of the reservoirs, it would be some sort of drain (e.g. an evil monster evaporated all of the water in the three reservoirs). The zero matrix acting on the reservoirs results in zero water left in each reservoir.

This does not obey water conservation — the total amount of water after the pump matrix is applied will not equal the initial total — but can still be represented as a matrix.

5.3.4 Conservation of Water

Now let's look at what happens when pumps move different amounts of water from each reservoir into other reservoirs. Specifically, let's work with this diagram:



Now, let us describe these pumps with this matrix:

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{1}{3} \\ 0 & \frac{3}{4} & \frac{1}{3} \end{bmatrix}$$

Each element of the matrix still represents a pump and indicates how much water is moved where. The first row indicates how much of each reservoir contributes to reservoir A when the pumps are activated. The second row does the same for reservoir B, and the third row is for reservoir C. For example, the upper left element $\frac{1}{2}$ tells us that half of what is in reservoir A will stay in reservoir A. Similarly, the $\frac{1}{2}$ on the middle left tells us that the other half of what was in reservoir A will flow into reservoir B when the pumps turn on. As we can see, each column of the matrix sums to one. This means the water is conserved (none is mysteriously lost or gained). The water will either stay in the original reservoir or move to a different one.

This is a useful fact about water moving between pools with no evaporation, but it is not something that is going to hold in all useful applications of matrices.

After activating the pumps once, how do we know how much water is in each reservoir? That can be calculated with a matrix-vector product, just as we saw with the simpler pump models:

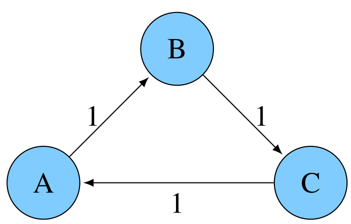
$$\begin{bmatrix} A_1 \\ B_1 \\ C_1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{1}{3} \\ 0 & \frac{3}{4} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} A_0 \\ B_0 \\ C_0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \cdot A_0 + \frac{1}{4} \cdot B_0 + \frac{1}{3} \cdot C_0 \\ \frac{1}{2} \cdot A_0 + 0 \cdot B_0 + \frac{1}{3} \cdot C_0 \\ 0 \cdot A_0 + \frac{3}{4} \cdot B_0 + \frac{1}{3} \cdot C_0 \end{bmatrix}$$

5.3.5 Matrix Multiplication Examples

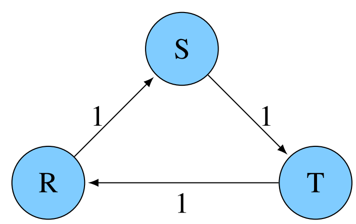
Now let's look at how matrix-matrix multiplication can be applied to water reservoirs and pumps.

5.3.6 Twin Cities

In this scenario, we have two cities that each have three reservoirs. The pump systems in the cities are identical. Let's start with the basic pump system.



(a) City 1



(b) City 2

The pump system can still be represented as a matrix, like before:

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

However, now the pump system acts on two sets of reservoirs instead of just one. Can our pump matrix act on two vectors representing the reservoirs instead of just one? We can combine the two vectors representing the water reservoirs in each city into a single matrix, with a column for each reservoir vector as follows:

$$\begin{bmatrix} A & R \\ B & S \\ C & T \end{bmatrix}$$

Now, let's use the pump matrix to find the water distribution in both cities in a single calculation. In City 1, the reservoirs initially have water amounts A_0, B_0, C_0. In City 2, the reservoirs initially have water amounts R_0, S_0, T_0. Once the pumps act on the reservoirs, the amount of water in each reservoir can be found through matrix-matrix multiplication:

$$\begin{bmatrix} A_1 & R_1 \\ B_1 & S_1 \\ C_1 & T_1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} A_0 & R_0 \\ B_0 & S_0 \\ C_0 & T_0 \end{bmatrix} = \begin{bmatrix} C_0 & T_0 \\ A_0 & R_0 \\ B_0 & S_0 \end{bmatrix}$$

If the cities have identical but more complicated pumps (such as the conservation pumps in the previous example), finding out how the reservoirs change is the same process. All that would be different is the "pump system" matrix.

What happens if you have the same pump-reservoir system in k cities? To find out how the pumps act on the reservoirs, you can still use matrix-matrix multiplication. One matrix describes the pumps, while the other describes the reservoirs. There would be k columns in the reservoir matrix, because each column is a vector that represents the reservoirs of a certain city.

In this case, we see a matrix acting on another matrix to transform multiple vectors the same way. Because of this, we can also see why the dimensions of the matrix have certain restrictions. The number of columns in the pumps matrix must match the number of rows in the reservoir matrix. The pumps matrix acts on each column of the reservoir matrix to produce a new column for the resulting matrix that describes amount of water for that city's reservoirs.

5.3.7 Activate Pumps Once... And Then Once More

Now, imagine we have one system of pumps for one city with three reservoirs. How can we calculate the amount of water in each reservoir after activating the pumps twice? From matrix-vector multiplication, we know how to find the amount after one activation. If matrix A represents the pumps and $\vec{v}_0$ represents the initial reservoir vector, then $\vec{v}_1 = A \cdot \vec{v}_0$ will tell us how much water is in each reservoir after one activation. Then $\vec{v}_2 = A \cdot \vec{v}_1$ will tell us how much water is in each reservoir after the second activation.

But from the reservoirs' standpoints, how they got from $\vec{v}_0$ to $\vec{v}_2$ does not matter. For all they know, it could have been some other system of pumps (matrix B) that acted on the initial reservoir vector ($\vec{v}_0$) that resulted in $\vec{v}_2$. This means that one set of pumps acting twice on the reservoirs is equivalent to *another* matrix acting on the reservoirs:

$$\begin{aligned} A\vec{v}_1 &= \vec{v}_2 \\ A(A \cdot \vec{v}_0) &= B \cdot \vec{v}_0 \\ (A \cdot A)\vec{v}_0 &= B \cdot \vec{v}_0 \\ A^2 &= B \end{aligned}$$

As an example, let's take the pump system from the conservation example in section 5.3.4:

$$A = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{1}{3} \\ 0 & \frac{3}{4} & \frac{1}{3} \end{bmatrix}, \vec{v}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Let's calculate $\vec{v_2}$:

$$\vec{v_1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{1}{3} \\ 0 & \frac{3}{4} & \frac{1}{3} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{13}{12} \\ \frac{5}{6} \\ \frac{13}{12} \end{bmatrix}$$

$$\vec{v_2} = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{1}{3} \\ 0 & \frac{3}{4} & \frac{1}{3} \end{bmatrix} \cdot \begin{bmatrix} \frac{13}{12} \\ \frac{5}{6} \\ \frac{13}{12} \end{bmatrix} = \begin{bmatrix} \frac{10}{9} \\ \frac{65}{72} \\ \frac{71}{72} \end{bmatrix}$$

For comparison:

$$B = A^2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{1}{3} \\ 0 & \frac{3}{4} & \frac{1}{3} \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{1}{3} \\ 0 & \frac{3}{4} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{3}{8} & \frac{3}{8} & \frac{13}{36} \\ \frac{1}{4} & \frac{3}{8} & \frac{5}{18} \\ \frac{3}{8} & \frac{1}{4} & \frac{13}{36} \end{bmatrix}$$

$$B \cdot \vec{v_0} = \begin{bmatrix} \frac{3}{8} & \frac{3}{8} & \frac{13}{36} \\ \frac{1}{4} & \frac{3}{8} & \frac{5}{18} \\ \frac{3}{8} & \frac{1}{4} & \frac{13}{36} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{10}{9} \\ \frac{65}{72} \\ \frac{71}{72} \end{bmatrix}$$

From this example, we can see that matrix-matrix multiplication results in an equivalent matrix. Pump system A acting twice on the reservoirs is the same as pump system B acting once on the reservoirs.

5.3.8 A Multitude of Pumps

Another example of matrix-matrix multiplication with these pumps and reservoirs is when two (or more) *different* sets of pumps act sequentially on a city's reservoirs. From the previous example, we know that a matrix multiplied by another matrix is equivalent to another matrix. That principle can be applied here. So if we have Pump System A act on the reservoirs ($\vec{v_0}$) and then Pump System B act on the reservoirs, it is the same as if some other Pump System C acted on the reservoirs:

$$\begin{aligned} B \cdot (A \cdot \vec{v_0}) &= C \cdot \vec{v_0} \\ (B \cdot A)\vec{v_0} &= C \cdot \vec{v_0} \\ B \cdot A &= C \end{aligned}$$

5.3.9 Continuous vs. Discrete Pumps

In all of the examples above, we've assumed that the pumps act instantaneously. That is, each time the pumps transfer water, first we calculate how much water will be moved based on the initial water levels in each reservoir. Then, all of the water is moved instantaneously. (At any given time, water entering reservoir A is not used in calculating how much water leaves reservoir A.) We can repeat this process (every day, minute, hour, etc), but each time all the water moves instantaneously. We call this process a *discrete time system* because water is transferred only at a discrete times.

While some physical properties happen in discrete time, others happen in *continuous time* – in other words, not instantaneously. We can describe continuous time systems with *differential equations*, which you will learn more about in EE 16B. But even without these techniques, we can still approximate the solution to a continuous time system by modeling it as a discrete time system where we take very small steps in time. Here, small is relative to how long the process takes: if it takes a minute for the pumps to transfer water, we could calculate the new water levels every second. If we are trying to model how light travels in space, we might need to calculate a new time step every femptosecond!

5.4 Practice Problems

These practice problems are also available in an interactive form on the course website.

1. Multiply $\begin{bmatrix} 1 & 5 & 0 \\ 10 & 3 & 7 \\ 6 & 4 & 11 \end{bmatrix}$ with $\begin{bmatrix} 2 & 12 & 3 \\ 1 & 8 & 0 \\ 9 & 1 & 2 \end{bmatrix}$. What is the first row of the resulting matrix?
   (a) $\begin{bmatrix} 16 & 52 & 5 \end{bmatrix}$
   (b) $\begin{bmatrix} 3 & 40 & 7 \end{bmatrix}$
   (c) $\begin{bmatrix} 7 & 52 & 3 \end{bmatrix}$
   (d) $\begin{bmatrix} 14 & 13 & 2 \end{bmatrix}$

2. **Matrix Multiplication**
   Consider the following matrices:

   $\mathbf{A_1} = \begin{bmatrix} 1 & 4 \end{bmatrix}$ $\mathbf{B_1} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ $\mathbf{A} = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$ $\mathbf{B} = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}$ $\mathbf{C} = \begin{bmatrix} 1 & 9 & 5 & 7 \\ 4 & 3 & 2 & 2 \end{bmatrix}$ $\mathbf{D} = \begin{bmatrix} 5 & 5 & 8 \\ 6 & 1 & 2 \\ 4 & 1 & 7 \\ 3 & 2 & 2 \end{bmatrix}$ $\mathbf{E} = \begin{bmatrix} 8 & 1 & 6 \\ 3 & 5 & 7 \\ 4 & 9 & 2 \end{bmatrix}$ $\mathbf{F} = \begin{bmatrix} 5 & 3 & 4 \\ 1 & 8 & 2 \\ 2 & 3 & 5 \end{bmatrix}$

   For each matrix multiplication problem, if the product exists, find the product by hand. Otherwise, explain why the product does not exist.

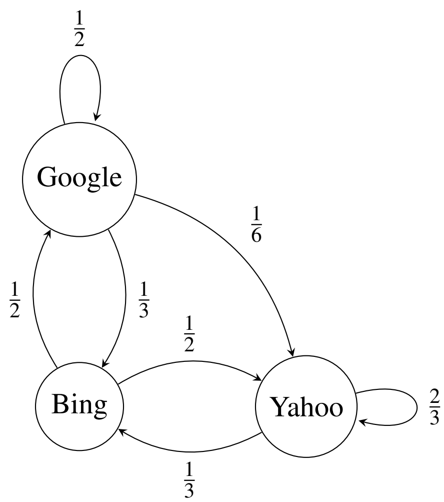
   (a) $\mathbf{A_1 B_1}$
   (b) **AB**
   (c) **BA**
   (d) **AC**
   (e) **DC**
   (f) **CD** (Write down the dimensions of the product if it exists. For practice, you can compute the product on your own)
   (g) **EF** (Practice on your own)
   (h) **FE** (Practice on your own)

3. Let the state transition matrix $\begin{bmatrix} 0.5 & 0.3 & 0 \\ 0 & 0.5 & 1 \\ 0.4 & 0.2 & 0 \end{bmatrix}$ represent people moving between three cities. If $\vec{x}[0] = \begin{bmatrix} 100 \\ 200 \\ 100 \end{bmatrix}$, find $\vec{x}[1]$.

4. Let the state transition matrix $\begin{bmatrix} 0.5 & 0.3 & 0 \\ 0 & 0.5 & 1 \\ 0.4 & 0.2 & 0 \end{bmatrix}$ represent people moving between three cities. Do people stay within these three cities?

5. Let the state transition matrix $\begin{bmatrix} 0.1 & 0.1 & 0.4 & 0.5 \\ 0.6 & 0.15 & 0 & 0.2 \\ 0.3 & 0.5 & 0.3 & 0.2 \\ 0 & 0.25 & 0.3 & 0.1 \end{bmatrix}$ represent the transfer of water between different buckets. The amount of water in each bucket a, b, c, and d at time n is $\begin{bmatrix} 3 \\ 4 \\ 19 \\ 1 \end{bmatrix}$. How much water is in bucket c at time $n+1$?

6. Consider the web traffic among the search engines given below. Write the state transition matrix for this system assuming that the state vector is $\vec{x}[t] = \begin{bmatrix} x_{\text{Google}}[t] \\ x_{\text{Yahoo}}[t] \\ x_{\text{Bing}}[t] \end{bmatrix}$.



7. Is the web traffic system modeled in the previous question conservative, i.e., is the number of web surfers in the system constant?

8. If a column adds up to a number larger than 1, what does this imply about the corresponding node?

   (a) People are leaving the system at that node.

   (b) People are entering the system at that node.

   (c) The node can exist in a conservative system.

   (d) The node has been wrongly modeled in the system.