

Main Goal: To construct a matrix that yields a desired transformation and to explore linear transformations through this example

By the end of the lesson students should be able to:

- Conceive of a matrix as a way of mapping elements of \mathbb{R}^2 to elements of \mathbb{R}^2
- Find a matrix that produces the desired geometric transformation from \mathbb{R}^2 to \mathbb{R}^2
- Determine whether the matrix for a transformation is unique (relative to the standard basis)
- Articulate criteria by which a set of vectors can be chosen to determine the (unique) transformation matrix (e.g., the set of vectors must span the domain)

The instructor should:

- NOT give away that the matrix is unique before students work on the task
- Define linear transformation
- Derive how AB can be interpreted as A acting on the columns of B after the students complete the task

Handout 1: Italicizing N

An image of Handout 1 is given below. See [page 48](#) for a printable student version of the handout.

THE ITALICIZING N PROBLEM

Name _____ Group Members _____

Suppose the "N" on the left is written in regular 12-point font. Find a matrix A that will transform N into the letter on the right, which is written in 'italics' in 16-point font.

$A =$

Work with a small group and write out your solution and approach. Make a list of any assumptions you notice your group making, or any questions for further pursuit that come to mind.

The task on Handout 1, which was inspired by an application problem in Lay's (2003) linear algebra textbook, is the experientially real starting point for this instructional sequence in that it draws on the experiences students have likely had observing font changes when working with computers. The task is deliberately placed on an unlabeled grid, pushing students to make their own choice about where to place the origin and how to coordinatize the space. In addition, it is deliberately left unclear what size the transformation matrix needs to be.

In this task, students must make choices of how to represent the "input" and "output" letter N using vectors, the dimensions of the matrix that will yield the desired transformation, and what set of input-output vectors to coordinate in order to find such a matrix. Students tend to develop systems of equations that can be pretty easily solved by inspection to find this matrix.

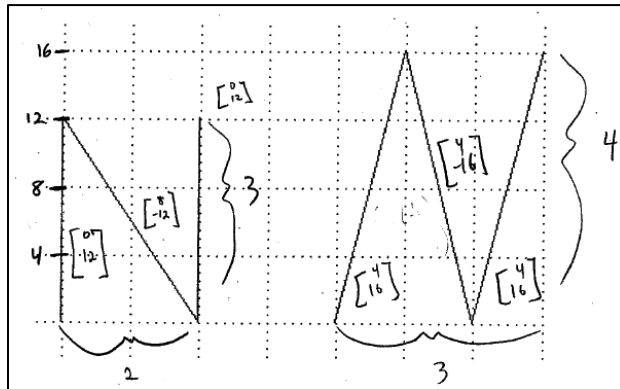
Overview for Task 1 of the Italicizing N Task Sequence

(Back to Top)

Handout 1: Italicizing N

Student Thinking

It is nontrivial for students to determine that both the inputs and outputs for the transformation lie in \mathbb{R}^2 and that a 2×2 matrix defines the mapping. Furthermore, students grapple with how to interpret and symbolize the representations of the N.

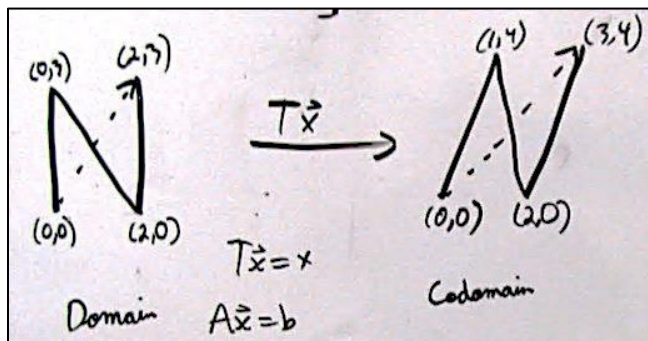
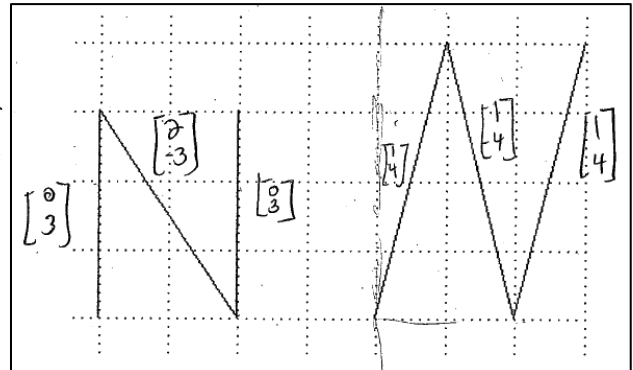


Group 1:

This group began by determining the scale of the x- and y-axes and naming each line segment of the letter N on the left. Notice that the left-most vertical segment is named $[0, 12]$, indicating that the group named this vector relative to an origin fixed on the bottom of this segment. However, the students named the second (diagonal) segment relative to an origin located at the upper left-hand corner of the letter N. Further, the last segment is named relative to an origin located on the bottom-right. The second N is labeled similarly, with the origin moving to the corners of N each time a new segment is named.

Group 2:

This group's approach was similar to group 1, with the exception that they used a different scale for the coordinates of the line segments. Notice again that, when naming each segment, the students varied the location of the origin and named each segment as though it were a vector itself.



Example 3

Unlike Groups 1 and 2, this group named points on N relative to a fixed origin at the lower left-hand corner of each N.

Regardless of the way in which students represent the letter N, a common approach is setting up a system of matrix equations – one matrix equation for each input-output pair – in order to determine the component values of A . Group 4 has shown their calculations for determining matrix A based on two input-output pairs. These pairs are consistent with Group 2's naming system for the line segments. Notice that the matrix A is initially written with parameters a , b , c , and d . Using this matrix, in chorus with the two input-output pairs, the group found two systems of two equations. The first system yielded constant values for b and d , which the group substituted into the second system of equations to find a and c . From this work, the group determined the matrix A .

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

$$A \begin{bmatrix} 2 \\ -3 \end{bmatrix} = \begin{bmatrix} 1 \\ -4 \end{bmatrix}$$

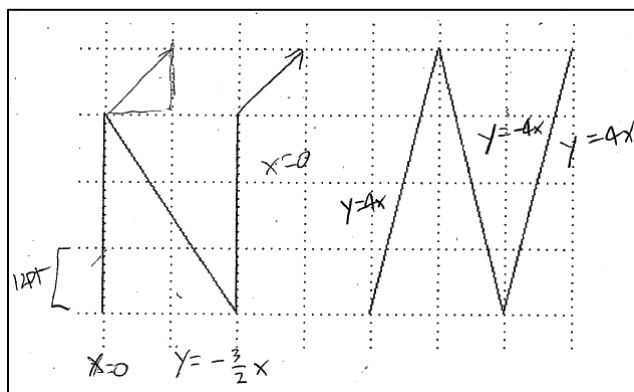
$$2 \begin{bmatrix} a \\ c \end{bmatrix} - 3 \begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} 1 \\ -4 \end{bmatrix}$$

$$\begin{cases} 2a - 3b = 1 \\ 2c - 3d = -4 \end{cases}$$

$$\begin{cases} 2a - 1 = 1 \\ 2a = 2 \\ a = 1 \end{cases}$$

$$\begin{cases} 2c - 3d = -4 \\ 2c - 3(\frac{1}{3}) = -4 \\ 2c - 4 = -4 \\ 2c = 0 \\ c = 0 \end{cases}$$

$$A = \begin{bmatrix} 1 & \frac{1}{3} \\ 0 & \frac{1}{3} \end{bmatrix}$$



This group attempted to determine the equation of the line that would be consistent with various parts of each N. Note that the equations corresponding to the vertical portions of the N on the left on the left are both $x = 0$. Similarly the equations of the parallel portions on the right are $y = 4x$. No further work with this approach was shown, and it is unclear how this work could help in correctly determining A .

This group's approach is similar to Group 4 in that they attempted to correspond input-output pairs with each other through matrix multiplication. Unlike Group 4, this group did not write A as a matrix of parameters. Also, this group concatenated the input and output vectors into 2×3 matrices. This could be viewed as an anticipation of matrix multiplication, seen as a repeated matrix multiplication on each of the three column vectors.

$$A = \begin{bmatrix} 0 & 2 & 0 \\ 3 & -3 & 3 \end{bmatrix} \quad N = \begin{bmatrix} 1 & 1 & 1 \\ 4 & -4 & 4 \end{bmatrix}$$

Getting Started with the Task

Following the Introduction mini-lecture ([see page 4](#)) on interpreting $A\mathbf{x} = \mathbf{b}$ as a transformation that acts on an input vector \mathbf{x} to produce the output vector \mathbf{b} , the instructor should pass out the Italicizing N task on Handout 1 ([see page 48](#) for the student printable version).

Again, take a moment to make sure everyone is “on the same page” of what the task is asking them to do, etc.

We recommend reading the task out loud with the students and giving them a few seconds to think over what is being asked of them before they get started on it.

We leave it to the instructor whether to discuss the location (and fixedness) of the origin before having the students engage in the Italicizing N task. This discussion would distinguish between approaches like those of Groups 1 and 2 from approaches like that of Group 3.

- One advantage of deciding (as a class) that the origin is at the bottom left corner of both Ns *before* the students engage in the task is that the students’ responses may be relatively more consistent, having selected a standard that everyone can follow. This approach, however, depending on the student population, can be seen as “doing” too much of the mathematics for the students.
- Allowing students to determine where the origin is for themselves allows them to engage in the task using their own naming system. This could develop into a debate in whole class discussion through which the students (along with the instructor) could navigate distinctions between various naming systems and democratically develop a convention for the class (that all points on the N will be thought of as the endpoint of a vector that emanates from a fixed origin). This discussion is nontrivial because the general mathematical community fixes the origin even though, in this case, both approaches are logically consistent and ultimately successful.

Working in Small Groups

This task involves a number of preliminary key steps for students to engage in during their small group work:

- Determining that \mathbb{R}^2 is the domain and the codomain for the transformation that A represents, so that means A will be a 2×2 matrix
- Determining a way to “coordinatize” the original N and the italicized N (e.g., the lower left corner of both Ns can be described by the point/vector $\langle 0,0 \rangle$ and the upper right corners of the Ns can be described by $\langle 2,3 \rangle$ and $\langle 3,4 \rangle$, respectively)
- Finding various input-output pairs that can be used to determine A

Additionally, the instructor should look out for more nuanced issues that students may encounter (explicitly or implicitly) that can be leveraged to help explore properties of transformation during whole class discussion. While these don’t need to be explored during small group work, the instructor may wish to make note of for whom these issues surfaced so as to facilitate a discussion of them during whole class discussion:

- How many input-output pairs are needed to determine the matrix A
- What type of input-output pairs are needed to determine the matrix A
- Whether the matrix A is unique

Deciding the dimension of A during Whole Class Discussion

Some classes may struggle with determining what the dimensions of A should be. If this is the case, it may be advantageous to stop the class partway through small group work to discuss that the two N s can be described as a collection of images in \mathbb{R}^2 , so A is taking input vectors from \mathbb{R}^2 (in the original N) and transforming them into output vectors in \mathbb{R}^2 (the 16-pt italicized N). Thus A has to be a 2×2 matrix. An example of how this conversation could go (inspired by how it has occurred in the past) is given:

Example Discussion

- Instructor:* So did anyone try to think about what the domain or codomain might be for these input and output space?
- Jonathan:* Yes. The regular N is your domain and then the italicized 16 N is your codomain.
- Instructor:* Okay, so you said the domain was the regular N , and the codomain was this italicized N . If we think back to the examples we saw before, we had spaces like \mathbb{R}^2 , \mathbb{R}^5 , \mathbb{R}^8 , etc. as our domains and codomains. So, here we could say that the original N is a collection of vectors in the domain, and the italicized N is a collection of vectors in the codomain. And if we think about the fact that the N and the italicized N are all lying on the plane and can be represented with 2 components, we'd say that the input space is \mathbb{R}^2 and the codomain is still \mathbb{R}^2 .

At this point, the instructor should give students more time in small group work to determine what the entries of A are. Note that the correct answer (according to the standard basis, which is left implied at the moment, is

$$A = \begin{bmatrix} 1 & 1/3 \\ 0 & 4/3 \end{bmatrix}.$$

Sharing solution approaches and the correct answer in Whole Class Discussion

If the class had not determined that all points on the N will be thought of as the endpoint of a vector that emanates from a fixed origin before they engaged in the task, examination of past student work has revealed two common notational strategies:

- Using vectors in \mathbb{R}^2 in which the 3 portions of the N are the vectors (e.g., Group 1-2)
- Using vectors in \mathbb{R}^2 (or points in the x - y plane) in which the N is represented with vectors whose endpoint lay on the N with tips originating from the same point on the letter (corresponding to a fixed origin) (e.g., Group 3)

To find the matrix A , students tend to either:

- Set up two matrix equations, such as $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$, convert them into two systems of equations, and solve for a , b , c , and d .
- Set up one matrix equation, such as $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 2 & 0 & 2 \\ 0 & 0 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 2 & 1 & 3 \\ 0 & 0 & 4 & 4 \end{bmatrix}$, convert it into a system of equations, and solve for a , b , c , and d .

The instructor could have students present their solution techniques and final solution for A to the class.

Defining “Linear Transformation” in whole class discussion

The problem in Task 1 that students have been working on is an example of a special type of transformation encountered in linear algebra, and so it is worth pausing to introduce and define that special class of transformations.

Two other reasons it is useful to define linear transformation here is (a) it can be used to show that any transformation that can be defined with a matrix is linear (and so we can use $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y}$ and $A(c\mathbf{x}) = cA\mathbf{x}$ in calculations and derivations; and (b) it is utilized in proofs of important results (such as knowing how a basis is transformed is sufficient to know how a whole space is transformed).

Definition: A transformation $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is called a **linear transformation** if it satisfies the following two properties: (1) $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$ for every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, and (2) $T(c\mathbf{x}) = cT(\mathbf{x})$ for every $\mathbf{x} \in \mathbb{R}^n$ and every $c \in \mathbb{R}$.

NOTE: This definition assumes \mathbb{R} as the field of scalars for the defined operations.

The instructor could point out that the students have seen examples of linear transformations, such as the derivative operator in Calculus: that is, $(f + g)' = f' + g'$.

A useful consequence of this definition is that the zero vector is always sent to the zero vector. Trying to prove why this would be true can be a good partner talk activity. Task 2 explores transformations of space at a more “global” level (rather than only one input-output pair at a time), so it is helpful to begin to hint that in a linear transformation, the zero vector is a sort of “anchor” that remains fixed and everything else is transformed around it.

Example Partner Talk Task:

Use the definition of linear transformation to prove that $T(0) = 0$. That is, that under a linear transformation, the zero vector in \mathbb{R}^n has to map to the zero vector in \mathbb{R}^m .

Possible justifications:

- Because T is linear, we know $T(cx) = cT(x)$ for every $x \in \mathbb{R}^n$ and every $c \in \mathbb{R}$. So choose $c = 0$ and consider what happens:

$$T(cx) = T(0x) = T(0) \quad \text{because the scalar 0 times the vector } x \text{ results in the zero vector in } \mathbb{R}^n \text{ for any vector } x \in \mathbb{R}^n$$

and

$$cT(x) = 0T(x) = 0 \quad \text{because } T(x) \text{ is a vector in } \mathbb{R}^m, \text{ and the scalar 0 times the vector } T(x) \text{ results in the zero vector in } \mathbb{R}^m$$

$$\therefore T(0) = 0.$$

- Because T is linear, we know $T(x + y) = T(x) + T(y)$ for every $x, y \in \mathbb{R}^n$. So choose the vector $y = -x$ and consider what happens:

$$T(x + y) = T(x + (-x)) = T(x - x) = T(0)$$

and

$$T(x) + T(y) = T(x) + T(-x) = T(x) - T(x) = 0$$

$$\therefore T(0) = 0.$$

It is beneficial to provide some time for students to explore examples that are and are not linear. This also involves needed practice in knowing what counts as evidence in either case (namely, that a transformation that is not linear needs to be shown to fail one of the two properties or $T(0) \neq 0$, whereas a transformation that is linear needs to be shown to satisfy both properties from the definition of linearity).

Below are a few possible examples. The instructor should choose which examples make sense for the class level, how many to do, how much time to spend in class discussing them, etc.

Example Discussion Questions:

Determine if the following transformations are linear or not:

- (a) $T: \mathbb{R} \rightarrow \mathbb{R}$ where $T(x) = x^2$
- (b) $T: \mathbb{R} \rightarrow \mathbb{R}$ where $T(x) = 4x$
- (c) $T: \mathbb{R} \rightarrow \mathbb{R}$ where $T(x) = x + 3$
- (d) $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ where $T(x_1, x_2) = (3x_1, x_1 + 2x_2)$
- (e) $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ where $T(x_1, x_2) = (x_1, 2x_2 - 1)$

(f) $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ where T is defined by $T(x) = Ax$, and $A = \begin{bmatrix} 1 & 1/3 \\ 0 & 4/3 \end{bmatrix}$.

(g) $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ where T is defined by $T(x) = Ax$, and $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $a, b, c, d \in \mathbb{R}$.

Of these examples, (a), (c), and (e) are not linear. Students most likely will be surprised that example (c) is not linear because the graph of $T(x) = x + 3$ “looks like a line.” Letting students discuss this example, and its similarity to example (e), is beneficial to their conceptual understanding of linear transformation.

Example (g) may want to be saved for homework. If the class is advanced and/or proof-oriented, the example “prove that any $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ that can be defined by an $m \times n$ matrix A is linear” may be given for homework of small group work.

Example (f) is the matrix that represents the “Italicizing N” transformation. That can serve as a nice example of how to use the definition to prove linearity:

$$\begin{aligned}
 \bullet \quad \begin{bmatrix} 1 & 1/3 \\ 0 & 4/3 \end{bmatrix} \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right) &= \begin{bmatrix} 1 & 1/3 \\ 0 & 4/3 \end{bmatrix} \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \end{bmatrix} = \begin{bmatrix} (x_1 + y_1) + \left(\frac{1}{3}\right)(x_2 + y_2) \\ \left(\frac{4}{3}\right)(x_2 + y_2) \end{bmatrix} \\
 &= \begin{bmatrix} x_1 + \left(\frac{1}{3}\right)x_2 + y_1 + \left(\frac{1}{3}\right)y_2 \\ \left(\frac{4}{3}\right)x_2 + \left(\frac{4}{3}\right)y_2 \end{bmatrix} \\
 &= \begin{bmatrix} x_1 + \left(\frac{1}{3}\right)x_2 \\ \left(\frac{4}{3}\right)x_2 \end{bmatrix} + \begin{bmatrix} y_1 + \left(\frac{1}{3}\right)y_2 \\ \left(\frac{4}{3}\right)y_2 \end{bmatrix} \\
 &= \begin{bmatrix} 1 & 1/3 \\ 0 & 4/3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 & 1/3 \\ 0 & 4/3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\
 &\quad \therefore T(x + y) = T(x) + T(y) \\
 \bullet \quad \begin{bmatrix} 1 & 1/3 \\ 0 & 4/3 \end{bmatrix} \left(c \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) &= \begin{bmatrix} 1 & 1/3 \\ 0 & 4/3 \end{bmatrix} \begin{bmatrix} cx_1 \\ cx_2 \end{bmatrix} = \begin{bmatrix} cx_1 + \left(\frac{1}{3}\right)cx_2 \\ \left(\frac{4}{3}\right)cx_2 \end{bmatrix} = c \begin{bmatrix} x_1 + \left(\frac{1}{3}\right)x_2 \\ \left(\frac{4}{3}\right)x_2 \end{bmatrix} = c \begin{bmatrix} 1 & 1/3 \\ 0 & 4/3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\
 &\quad \therefore T(cx) = cT(x) \\
 \bullet \quad &\text{The two properties of a linear transformation (distribution across vector addition and scalar multiplication) could be illustrated through considering specific input-output pairs of the Italicizing N transformation:} \\
 &\quad \circ \text{ Choosing the bottom left corner of the N to be the origin (in both Ns) illustrates that zero gets sent to zero} \\
 &\quad \circ \text{ One could consider the “parallelogram rule,” that } \begin{bmatrix} 2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \text{ and } T\left(\begin{bmatrix} 2 \\ 3 \end{bmatrix}\right) = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \\
 &\quad \text{or that } T\left(\begin{bmatrix} 2 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, T\left(\begin{bmatrix} 0 \\ 3 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \text{ and } \begin{bmatrix} 2 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}.
 \end{aligned}$$

Exploring Nuanced Issues during Whole Class Discussion

The instructor may have noticed nuanced issues that students encountered (explicitly or implicitly) during small group work. There is no linear order in which these need to be discussed, and it is up to the instructor to decide the rigor with which to discuss these issues, if at all.

- How many input-output pairs are needed to determine the matrix A
- What type of input-output pairs are needed to determine the matrix A
- Whether the matrix A is unique

The responses to these three questions are somewhat bound together. For instance, if students did not choose input vectors that spanned \mathbb{R}^2 in order to determine A , the matrix A they found may not be unique. However, that

matrix A would not perform the desired italicizing transformation for all possible input vectors. Thus, to find an A that works for all inputs as desired, one needs to choose input vectors that span the domain. If this occurs, the matrix A will be the unique matrix for that transformation (up to a choice of basis).

Additionally, the teacher may be able to prompt students to consider what the fewest number of input-output pairs needed is so that A can be determined. As such, if the term **basis** has not already been defined in the linear algebra course, this may be a natural place to do so:

Definition: A **basis** for a vector space V is a set of linearly independent vectors in V that spans V .

NOTE: An instructor may choose to use a less general definition, such as “A basis for \mathbb{R}^n is a set of linearly independent vectors in \mathbb{R}^n that spans \mathbb{R}^n .”

- Thus, the class may come to establish that, in order to determine the $m \times n$ matrix A for a transformation, they only need to know the input-output pairs for inputs that form a basis for \mathbb{R}^n .
- This can be leveraged to explain that if one knows what happens to a basis under a transformation, then one knows what happens to any vector in the domain:

(Assuming $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$) Suppose \mathbf{v} and \mathbf{w} are chosen as input vectors, \mathbf{v} and \mathbf{w} form a basis for \mathbb{R}^2 , and how \mathbf{v} and \mathbf{w} are transformed under T is known. Because they form a basis, any vector in \mathbb{R}^2 can be written as a linear combination of \mathbf{v} and \mathbf{w} . That is, for any \mathbf{u} in \mathbb{R}^2 there exists scalars a and b such that $\mathbf{u} = a\mathbf{v} + b\mathbf{w}$. Next consider $T(\mathbf{u})$. $T(\mathbf{u}) = T(a\mathbf{v} + b\mathbf{w}) = T(a\mathbf{v}) + T(b\mathbf{w})$ by linearity of T , which equals $aT(\mathbf{v}) + bT(\mathbf{w})$, also by linearity of T . Thus $T(\mathbf{u}) = aT(\mathbf{v}) + bT(\mathbf{w})$ for any \mathbf{u} . Thus, because it is known what happens to \mathbf{v} and \mathbf{w} under T , what happens to any \mathbf{u} under T can also be determined.

- The teacher may wish to further connect to **basis** by discussing how the columns of A are the images of the (ordered) basis vectors. Here, where the standard basis has been assumed, the first column of A is the image of \mathbf{e}_1 (which is just \mathbf{e}_1), and the second column of A is the image of \mathbf{e}_2 (which is $\langle 1/3, 4/3 \rangle$). While this instructional unit may be completed without this information, it may be nice to say it here to connect with later transformations within eigen theory and change of basis.

Exploring a matrix times a matrix in whole class discussion

As mentioned previously, some students might try to find $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ by “concatenating” multiple input-output pairs and then make use of a system of equations to solve for A , such as:

- $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & 4 \end{bmatrix}$ (use 2 input-outputs pairs from the N)
- $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 2 & 0 & 2 & 0 \\ 0 & 3 & 3 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 3 & 0 \\ 0 & 4 & 4 & 0 \end{bmatrix}$ (use 4 input-outputs pairs from the N)

However, to this point, A acting on inputs to produce outputs has only been defined for one pair at a time: $A\mathbf{x} = \mathbf{b}$. Why this “concatenation” is valid is a great way to develop that AB can be defined as A acting on the columns of B :

Example Discussion Question:

Why is it ok to use more than one input-output pair at a time to determine A ?

In other words, we know $A\mathbf{x}_1 = \mathbf{b}_1$ is valid, but is $A[\mathbf{x}_1 \quad \mathbf{x}_2 \quad \dots \quad \mathbf{x}_n] = [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \dots \quad \mathbf{b}_n]$ ok?

One way to explore this is to consider $[\mathbf{x}_1 \quad \mathbf{x}_2 \quad \dots \quad \mathbf{x}_n]$ as a matrix. Let’s call it D . So we’re trying to determine how to interpret / compute $AD = A[\mathbf{x}_1 \quad \mathbf{x}_2 \quad \dots \quad \mathbf{x}_n]$

To help out, let's consider $AD\mathbf{y}$ for some vector $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$.

$$\begin{aligned} \text{So } AD\mathbf{y} &= A(D\mathbf{y}) = A \left(\begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \right) \\ &= A(y_1\mathbf{x}_1 + y_2\mathbf{x}_2 + \dots + y_n\mathbf{x}_n) \\ &= Ay_1\mathbf{x}_1 + Ay_2\mathbf{x}_2 + \dots + Ay_n\mathbf{x}_n \text{ because of linearity} \\ &= y_1A\mathbf{x}_1 + y_2A\mathbf{x}_2 + \dots + y_nA\mathbf{x}_n \text{ because of linearity} \\ &= \begin{bmatrix} A\mathbf{x}_1 & A\mathbf{x}_2 & \dots & A\mathbf{x}_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \\ &= \begin{bmatrix} A\mathbf{x}_1 & A\mathbf{x}_2 & \dots & A\mathbf{x}_n \end{bmatrix} \mathbf{y} \end{aligned}$$

$$\therefore AD = A \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \end{bmatrix} = \begin{bmatrix} A\mathbf{x}_1 & A\mathbf{x}_2 & \dots & A\mathbf{x}_n \end{bmatrix}$$

Therefore, the matrix multiplication AD can be interpreted as A acting on the columns of D , and student approaches such as $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 2 & 0 & 2 & 0 \\ 0 & 3 & 3 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 3 & 0 \\ 0 & 4 & 4 & 0 \end{bmatrix}$ are valid.

An instructor could leverage this conversation to look at what matrices can be multiplied together and what dimensions the resulting matrix would be:

Example Discussion Question:

So, in general, what dimensions do the matrices A and D have to be so that the multiplication is defined, and what dimensions would the resulting matrix F be in $AD = F$?

$$\begin{array}{ccc} \begin{array}{c} n \text{ columns} \\ m \text{ rows} \end{array} \left[\begin{array}{c} \phantom{\hspace{1cm}} \\ \phantom{\hspace{1cm}} \\ \phantom{\hspace{1cm}} \end{array} \right] & \begin{array}{c} \text{Can have any number of columns;} \\ \text{say it has } p \text{ columns} \\ \left[\begin{array}{c} \phantom{\hspace{1cm}} \\ \phantom{\hspace{1cm}} \\ \phantom{\hspace{1cm}} \end{array} \right] \\ \text{Needs } n \text{ rows so that} \\ \text{there are exactly} \\ \text{enough components to} \\ \text{be weights for the } n \\ \text{columns of } A \end{array} & = & \left[\begin{array}{c} \phantom{\hspace{1cm}} \\ \phantom{\hspace{1cm}} \\ \phantom{\hspace{1cm}} \end{array} \right] \begin{array}{c} F \text{ has } p \text{ columns} \\ \text{because } D \text{ has } p \\ \text{columns, and } F \text{ is} \\ \text{the result of } A \\ \text{acting on the } p \\ \text{columns of } D. \\ \text{For each column } \mathbf{f}_i, A\mathbf{x}_i = \mathbf{f}_i \\ \text{because it's the result of } A \\ \text{acting on the columns of } D. \\ \text{The columns of } D \text{ contained} \\ m\text{-components, so the} \\ \text{columns of } F \text{ do, too. Thus } F \\ \text{has } m \text{ rows.} \end{array} \end{array}$$

Leads to the commonly known "trick":

$$(m \times n)(n \times p) = (m \times p)$$