52. $T(\vec{x}) = \begin{bmatrix} 0 & 4 \\ 5 & -3 \end{bmatrix} \vec{x}$

53. Sketch the image of the unit circle under the linear transformation

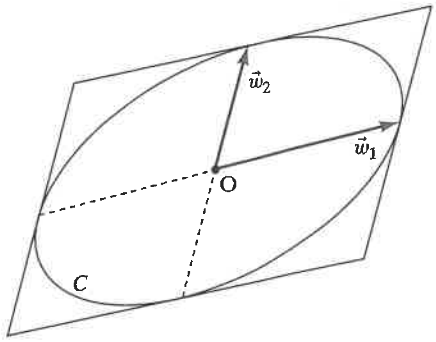
$$T(\vec{x}) = \begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix} \vec{x}.$$

54. Let T be an invertible linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$. Show that the image of the unit circle is an ellipse centered at the origin.[8] *Hint*: Consider two perpendicular unit vectors $\vec{v}_1$ and $\vec{v}_2$ such that $T(\vec{v}_1)$ and $T(\vec{v}_2)$ are perpendicular. See Exercise 47. The unit circle consists of all vectors of the form

$$\vec{v} = \cos(t)\vec{v}_1 + \sin(t)\vec{v}_2,$$

where t is a parameter.

55. Let $\vec{w}_1$ and $\vec{w}_2$ be two nonparallel vectors in $\mathbb{R}^2$. Consider the curve C in $\mathbb{R}^2$ that consists of all vectors of the form $\cos(t)\vec{w}_1 + \sin(t)\vec{w}_2$, where t is a parameter. Show that C is an ellipse. *Hint:* You can interpret C as the image of the unit circle under a suitable linear transformation; then use Exercise 54.



56. Consider an invertible linear transformation T from $\mathbb{R}^2$ to $\mathbb{R}^2$. Let C be an ellipse in $\mathbb{R}^2$. Show that the image of C under T is an ellipse as well. *Hint*: Use the result of Exercise 55.

2.3 Matrix Products

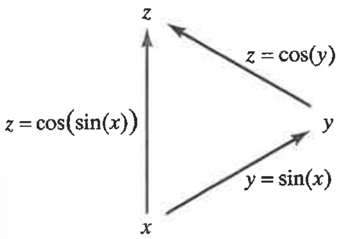


Figure 1

Recall the *composition* of two functions: The composite of the functions $y = \sin(x)$ and $z = \cos(y)$ is $z = \cos(\sin(x))$, as illustrated in Figure 1.

Similarly, we can compose two linear transformations.

To understand this concept, let's return to the coding example discussed in Section 2.1. Recall that the position $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ of your boat is encoded and that you radio the encoded position $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ to Marseille. The coding transformation is

$$\vec{y} = A\vec{x}, \quad \text{with} \quad A = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}.$$

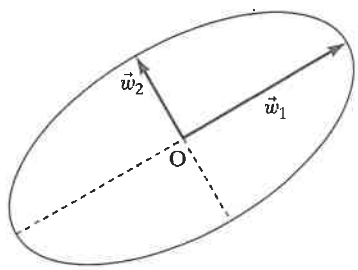
In Section 2.1, we left out one detail: Your position is radioed on to Paris, as you would expect in a centrally governed country such as France. Before broadcasting to Paris, the position $\vec{y}$ is again encoded, using the linear transformation

[8] An ellipse in $\mathbb{R}^2$ centered at the origin may be defined as a curve that can be parametrized as

$$\cos(t)\vec{w}_1 + \sin(t)\vec{w}_2,$$

for two perpendicular vectors $\vec{w}_1$ and $\vec{w}_2$. Suppose the length of $\vec{w}_1$ exceeds the length of $\vec{w}_2$. Then we call the vectors $\pm\vec{w}_1$ the semimajor axes of the ellipse and $\pm\vec{w}_2$ the semiminor axes.

Convention: All ellipses considered in this text are centered at the origin unless stated otherwise.



$$\vec{z} = B\vec{y}, \quad \text{with} \quad B = \begin{bmatrix} 6 & 7 \\ 8 & 9 \end{bmatrix}$$

this time, and the sailor in Marseille radios the encoded position $\vec{z}$ to
Figure 2.

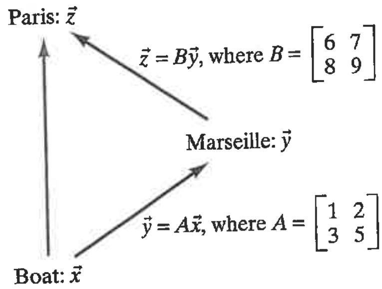


Figure 2

We can think of the message $\vec{z}$ received in Paris as a function of
position $\vec{x}$ of the boat,

$$\vec{z} = B(A\vec{x}),$$

the composite of the two transformations $\vec{y} = A\vec{x}$ and $\vec{z} = B\vec{y}$. Is this t
tion $\vec{z} = T(\vec{x})$ linear, and, if so, what is its matrix? We will show two ap
these important questions: (a) using brute force, and (b) using some the

a. We write the components of the two transformations and substit

$$\begin{aligned} z_1 &= 6y_1 + 7y_2 \\ z_2 &= 8y_1 + 9y_2 \end{aligned} \quad \text{and} \quad \begin{aligned} y_1 &= x_1 + 2x_2 \\ y_2 &= 3x_1 + 5x_2 \end{aligned}$$

so that

$$\begin{aligned} z_1 &= 6(x_1 + 2x_2) + 7(3x_1 + 5x_2) = (6 \cdot 1 + 7 \cdot 3)x_1 + (6 \cdot 2 \\ &= 27x_1 + 47x_2, \\ z_2 &= 8(x_1 + 2x_2) + 9(3x_1 + 5x_2) = (8 \cdot 1 + 9 \cdot 3)x_1 + (8 \cdot 2 \\ &= 35x_1 + 61x_2. \end{aligned}$$

This shows that the composite is indeed linear, with matrix

$$\begin{bmatrix} 6 \cdot 1 + 7 \cdot 3 & 6 \cdot 2 + 7 \cdot 5 \\ 8 \cdot 1 + 9 \cdot 3 & 8 \cdot 2 + 9 \cdot 5 \end{bmatrix} = \begin{bmatrix} 27 & 47 \\ 35 & 61 \end{bmatrix}.$$

b. We can use Theorem 1.3.10 to show that the transformation $T($
is linear:

$$\begin{aligned} T(\vec{v} + \vec{w}) &= B\big(A(\vec{v} + \vec{w})\big) = B(A\vec{v} + A\vec{w}) \\ &= B(A\vec{v}) + B(A\vec{w}) = T(\vec{v}) + T(\vec{w}), \\ T(k\vec{v}) &= B\big(A(k\vec{v})\big) = B\big(k(A\vec{v})\big) = k\big(B(A\vec{v})\big) = k \end{aligned}$$

Once we know that T is linear, we can find its matrix by c
vectors $T(\vec{e}_1) = B(A\vec{e}_1)$ and $T(\vec{e}_2) = B(A\vec{e}_2)$; the matrix
$\begin{bmatrix} T(\vec{e}_1) & T(\vec{e}_2) \end{bmatrix}$, by Theorem 2.1.2:

$$T(\vec{e}_1) = B(A\vec{e}_1) = B(\text{first column of } A) = \begin{bmatrix} 6 & 7 \\ 8 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} =$$

$$T(\vec{e}_2) = B(A\vec{e}_2) = B(\text{second column of } A) = \begin{bmatrix} 6 & 7 \\ 8 & 9 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

We find that the matrix of the linear transformation $T(\vec{x}) = B(A\vec{x})$ is

$$\begin{bmatrix} | & | \\ T(\vec{e}_1) & T(\vec{e}_2) \\ | & | \end{bmatrix} = \begin{bmatrix} 27 & 47 \\ 35 & 61 \end{bmatrix}.$$

This result agrees with the result in (a), of course.

The matrix of the linear transformation $T(\vec{x}) = B(A\vec{x})$ is called the *product* of the matrices B and A, written as BA. This means that

$$T(\vec{x}) = B(A\vec{x}) = (BA)\vec{x},$$

for all vectors $\vec{x}$ in $\mathbb{R}^2$. See Figure 3.

Now let's look at the product of larger matrices. Let B be an $n \times p$ matrix and A a $p \times m$ matrix. These matrices represent linear transformations, as shown in Figure 4.

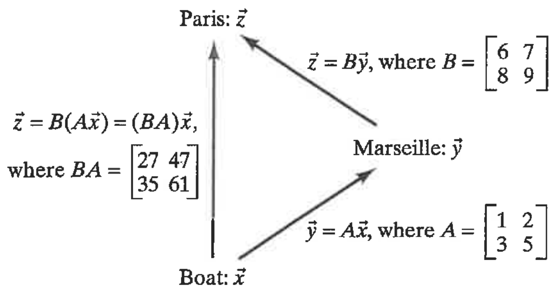


Figure 3

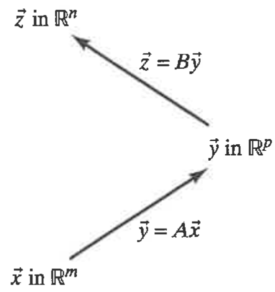


Figure 4

Again, the composite transformation $\vec{z} = B(A\vec{x})$ is linear. [Part (b) of the foregoing justification applies in this more general case as well.] The matrix of the linear transformation $\vec{z} = B(A\vec{x})$ is called the *product* of the matrices B and A, written as BA. Note that BA is an $n \times m$ matrix (as it represents a linear transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$). As in the case of $\mathbb{R}^2$, the equation

$$\vec{z} = B(A\vec{x}) = (BA)\vec{x}$$

holds for all vectors $\vec{x}$ in $\mathbb{R}^m$, by definition of the product BA. See Figure 5.

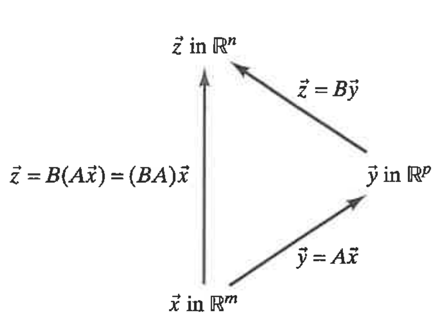


Figure 5

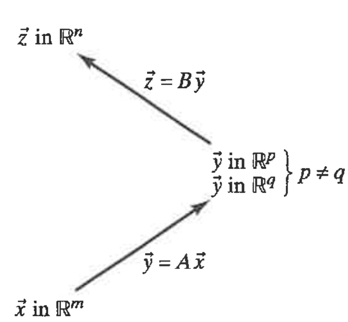


Figure 6

In the definition of the matrix product BA, the number of columns of B matches the number of rows of A. What happens if these two numbers are different? Suppose B is an $n \times p$ matrix and A is a $q \times m$ matrix, with $p \neq q$.

In this case, the transformations $\vec{z} = B\vec{y}$ and $\vec{y} = A\vec{x}$ cannot be composed, since the target space of $\vec{y} = A\vec{x}$ is different from the domain of $\vec{z} = B\vec{y}$. See

Figure 6. To put it more plainly: The output of $\vec{y} = A\vec{x}$ is not an accept
for the transformation $\vec{z} = B\vec{y}$. In this case, the matrix product BA is unc

Definition 2.3.1 **Matrix multiplication**

a. Let B be an $n \times p$ matrix and A a $q \times m$ matrix. The produ
defined if (and only if) $p = q$.

b. If B is an $n \times p$ matrix and A a $p \times m$ matrix, then the produ
defined as the matrix of the linear transformation $T(\vec{x}) = B($
means that $T(\vec{x}) = B(A\vec{x}) = (BA)\vec{x}$, for all $\vec{x}$ in the vector s
The product BA is an $n \times m$ matrix.

Although this definition of matrix multiplication does not give u
instructions for computing the product of two numerically given mat
instructions can be derived easily from the definition.

As in Definition 2.3.1, let B be an $n \times p$ matrix and A a $p \times m$ m
think about the columns of the matrix BA:

$$\begin{aligned}(i\text{th column of } BA) &= (BA)\vec{e}_i \\ &= B(A\vec{e}_i) \\ &= B(i\text{th column of } A).\end{aligned}$$

If we denote the columns of A by $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$, we can write

$$BA = B\begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_m \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ B\vec{v}_1 & B\vec{v}_2 & \cdots & B\vec{v}_m \\ | & | & & | \end{bmatrix}$$

Theorem 2.3.2 **The columns of the matrix product**

Let B be an $n \times p$ matrix and A a $p \times m$ matrix with columns $\vec{v}_1$,
Then, the product BA is

$$BA = B\begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_m \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ B\vec{v}_1 & B\vec{v}_2 & \cdots & B\vec{v}_m \\ | & | & & | \end{bmatrix}$$

To find BA, we can multiply B by the columns of A and c
resulting vectors.

This is exactly how we computed the product

$$BA = \begin{bmatrix} 6 & 7 \\ 8 & 9 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 27 & 47 \\ 35 & 61 \end{bmatrix}$$

on page 76, using approach (b).

For practice, let us multiply the same matrices in the reverse o
column of AB is $\begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} 6 \\ 8 \end{bmatrix} = \begin{bmatrix} 22 \\ 58 \end{bmatrix}$; the second is $\begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} 7 \\ 9 \end{bmatrix} =$

$$AB = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} 6 & 7 \\ 8 & 9 \end{bmatrix} = \begin{bmatrix} 22 & 25 \\ 58 & 66 \end{bmatrix}.$$

Compare the two previous displays to see that $AB \neq BA$: Mat
tion is *noncommutative*. This should come as no surprise, in view

the matrix product represents a composite of transformations. Even for functions of one variable, the order in which we compose matters. Refer to the first example in this section and note that the functions $\cos(\sin(x))$ and $\sin(\cos(x))$ are different.

Theorem 2.3.3

> **Matrix multiplication is noncommutative**
>
> $AB \neq BA$, in general. However, at times it does happen that $AB = BA$; then we say that the matrices A and B *commute*.

It is useful to have a formula for the ijth entry of the product BA of an $n \times p$ matrix B and a $p \times m$ matrix A.

Let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$ be the columns of A. Then, by Theorem 2.3.2,

$$BA = B\begin{bmatrix} | & | & & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_j & \cdots & \vec{v}_m \\ | & | & & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | & & | \\ B\vec{v}_1 & B\vec{v}_2 & \cdots & B\vec{v}_j & \cdots & B\vec{v}_m \\ | & | & & | & & | \end{bmatrix}.$$

The ijth entry of the product BA is the ith component of the vector $B\vec{v}_j$, which is the dot product of the ith row of B and $\vec{v}_j$, by Definition 1.3.7.

Theorem 2.3.4

> **The entries of the matrix product**
>
> Let B be an $n \times p$ matrix and A a $p \times m$ matrix. The ijth entry of BA is the dot product of the ith row of B with the jth column of A.
>
> $$BA = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{ip} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pj} & \cdots & a_{pm} \end{bmatrix}$$
>
> is the $n \times m$ matrix whose ijth entry is
>
> $$b_{i1}a_{1j} + b_{i2}a_{2j} + \cdots + b_{ip}a_{pj} = \sum_{k=1}^{p} b_{ik}a_{kj}.$$

EXAMPLE 1

$$\begin{bmatrix} 6 & 7 \\ 8 & 9 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 6\cdot 1 + 7\cdot 3 & 6\cdot 2 + 7\cdot 5 \\ 8\cdot 1 + 9\cdot 3 & 8\cdot 2 + 9\cdot 5 \end{bmatrix} = \begin{bmatrix} 27 & 47 \\ 35 & 61 \end{bmatrix}$$

We have done these computations before. (Where?) ■

EXAMPLE 2 Compute the products BA and AB for $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$. Interpret your answers geometrically, as composites of linear transformation. Draw composition diagrams.

Solution

$$BA = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad AB = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Note that in this special example it turns out that $BA = -AB$.

From Section 2.2 we recall the following geometrical interpretations:

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ represents the reflection about the vector } \begin{bmatrix} 1 \\ 1 \end{bmatrix};$$

$$B = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \text{ represents the reflection about } \begin{bmatrix} 0 \\ 1 \end{bmatrix};$$

$$BA = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \text{ represents the rotation through } \frac{\pi}{2}; \text{ and}$$

$$AB = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \text{ represents the rotation through } -\frac{\pi}{2}.$$

Let's use our standard L to show the effect of these transformations. See and 8.

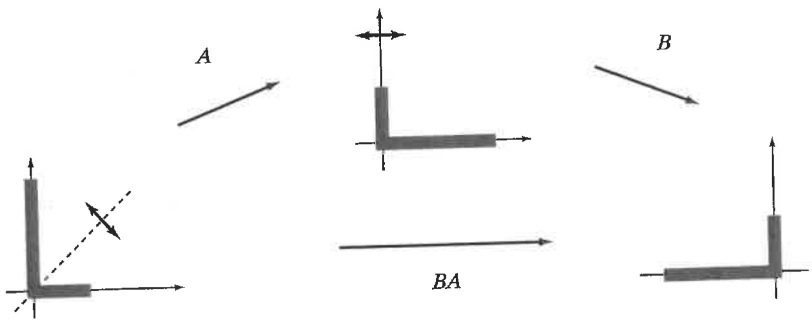


Figure 7

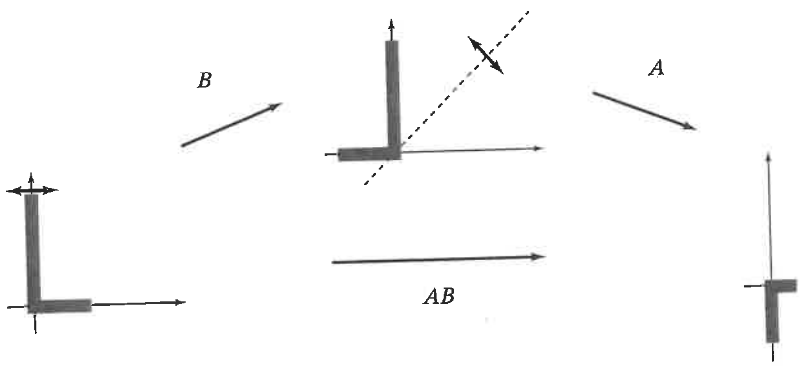


Figure 8

Matrix Algebra

Next let's discuss some algebraic rules for matrix multiplication.

- Composing a linear transformation with the identity transformati side, leaves the transformation unchanged. See Example 2.1.4.

Theorem 2.3.5

> **Multiplying with the identity matrix**
>
> For an $n \times m$ matrix A,
>
> $$AI_m = I_n A = A.$$

- If A is an $n \times p$ matrix, B a $p \times q$ matrix, and C a $q \times m$ matrix, what is the relationship between $(AB)C$ and $A(BC)$?

  One way to think about this problem (although perhaps not the most elegant one) is to write C in terms of its columns: $C = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_m \end{bmatrix}$. Then

$$(AB)C = (AB)\begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_m \end{bmatrix} = \begin{bmatrix} (AB)\vec{v}_1 & (AB)\vec{v}_2 & \cdots & (AB)\vec{v}_m \end{bmatrix},$$

and

$$A(BC) = A\begin{bmatrix} B\vec{v}_1 & B\vec{v}_2 & \cdots & B\vec{v}_m \end{bmatrix} = \begin{bmatrix} A(B\vec{v}_1) & A(B\vec{v}_2) & \cdots & A(B\vec{v}_m) \end{bmatrix}.$$

Since $(AB)\vec{v}_i = A(B\vec{v}_i)$, by definition of the matrix product, we find that $(AB)C = A(BC)$.

Theorem 2.3.6 **Matrix multiplication is associative**

$$(AB)C = A(BC)$$

We can simply write ABC for the product $(AB)C = A(BC)$.

A more conceptual proof is based on the fact that the composition of functions is associative. The two linear transformations

$$T(\vec{x}) = \big((AB)C\big)\vec{x} \quad \text{and} \quad L(\vec{x}) = \big(A(BC)\big)\vec{x}$$

are identical because, by the definition of matrix multiplication,

$$T(\vec{x}) = \big((AB)C\big)\vec{x} = (AB)(C\vec{x}) = A\big(B(C\vec{x})\big)$$

and

$$L(\vec{x}) = \big(A(BC)\big)\vec{x} = A\big((BC)\vec{x}\big) = A\big(B(C\vec{x})\big).$$

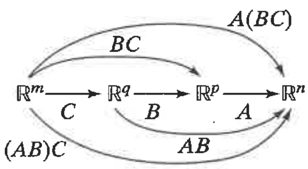


Figure 9

The domains and target spaces of the linear transformations defined by the matrices A, B, C, BC, AB, $A(BC)$, and $(AB)C$ are shown in Figure 9.

Theorem 2.3.7 **Distributive property for matrices**

If A and B are $n \times p$ matrices, and C and D are $p \times m$ matrices, then

$$A(C + D) = AC + AD, \quad \text{and}$$
$$(A + B)C = AC + BC.$$

You will be asked to verify this property in Exercise 27.

Theorem 2.3.8 If A is an $n \times p$ matrix, B is a $p \times m$ matrix, and k is a scalar, then

$$(kA)B = A(kB) = k(AB).$$

You will be asked to verify this property in Exercise 28.

Block Matrices (Optional)

In the popular puzzle Sudoku, one considers a 9×9 matrix A that is subdivided into nine 3×3 matrices called *blocks*. The puzzle setter provides some of the 81 entries of matrix A, and the objective is to fill in the remaining entries so that each row of A, each column of A, and each block contains each of the digits 1 through 9 exactly once.

5	3			7				
6			1	9	5			
	9	8					6	
8				6				3
4			8		3			1
7				2				6
	6					2	8	
			4	1	9			5
				8			7	9

This Sudoku puzzle is an example of a *block matrix* (or *partitioned matri* a matrix that is partitioned into rectangular submatrices, called blocks, by horizontal and vertical lines that go all the way through the matrix.

The blocks need not be of equal size.

For example, we can partition the matrix

$$B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 6 & 7 & 9 \end{bmatrix} \quad \text{as} \quad B = \left[\begin{array}{cc|c} 1 & 2 & 3 \\ 4 & 5 & 6 \\ \hline 6 & 7 & 9 \end{array}\right] = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

where $B_{11} = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}$, $B_{12} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$, $B_{21} = \begin{bmatrix} 6 & 7 \end{bmatrix}$, and $B_{22} = [9]$.

A useful property of block matrices is the following:

Theorem 2.3.9 **Multiplying block matrices**

Block matrices can be multiplied as though the blocks were scalars (i the formula in Theorem 2.3.4):

$$AB = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1p} \\ A_{21} & A_{22} & \cdots & A_{2p} \\ \vdots & \vdots & & \vdots \\ A_{i1} & A_{i2} & \cdots & A_{ip} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{np} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1j} & \cdots \\ B_{21} & B_{22} & \cdots & B_{2j} & \cdots \\ \vdots & \vdots & & \vdots & \\ B_{p1} & B_{p2} & \cdots & B_{pj} & \cdots \end{bmatrix}$$

is the block matrix whose ijth block is the matrix

$$A_{i1}B_{1j} + A_{i2}B_{2j} + \cdots + A_{ip}B_{pj} = \sum_{k=1}^{p} A_{ik}B_{kj},$$

provided that all the products $A_{ik}B_{kj}$ are defined.

Verifying this fact is left as an exercise. A numerical example foll

EXAMPLE 3 $$\left[\begin{array}{cc|c} 0 & 1 & -1 \\ 1 & 0 & 1 \end{array}\right] \left[\begin{array}{cc|c} 1 & 2 & 3 \\ 4 & 5 & 6 \\ \hline 7 & 8 & 9 \end{array}\right]$$

$$= \left[\left[\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}\right]\left[\begin{matrix} 1 & 2 \\ 4 & 5 \end{matrix}\right] + \left[\begin{matrix} -1 \\ 1 \end{matrix}\right]\left[\begin{matrix} 7 & 8 \end{matrix}\right] \;\middle|\; \left[\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}\right]\left[\begin{matrix} 3 \\ 6 \end{matrix}\right] + \left[\begin{matrix} -1 \\ 1 \end{matrix}\right][9\right.$$

$$= \left[\begin{array}{cc|c} -3 & -3 & -3 \\ 8 & 10 & 12 \end{array}\right].$$

Compute this product without using a partition, and see whether you find the same result. ■

In this simple example, using blocks is somewhat pointless. Example 3 merely illustrates Theorem 2.3.9. In Example 2.4.7, we will see a more sensible usage of the concept of block matrices.

Powers of Transition Matrices

We will conclude this section with an example on transition matrices. See Definition 2.1.4.

EXAMPLE 4 Let's revisit the mini-Web we considered in Example 9 of Section 2.1:

$$\begin{array}{ccc} 1 & \rightleftarrows & 2 \\ \downarrow & \swarrow & \uparrow \\ 3 & \rightarrow & 4 \end{array}$$

with the transition matrix

$$A = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

At a predetermined time, all the surfers will follow one of the available links, in the manner described in Example 2.1.9. If the initial distribution of the surfers among the four pages is given by the vector $\vec{x}$, then the distribution after this transition will be $A\vec{x}$. Now, let's iterate this process: Imagine an event of "speed surfing," where, every few minutes, at the blow of a whistle, each surfer will follow an available link. After two transitions, the distribution will be $A(A\vec{x}) = A^2\vec{x}$, and after m transitions the distribution will be given by the vector $A^m\vec{x}$. Let's use technology to compute some of the powers A^m of matrix A:

$$A^2 = \begin{bmatrix} \frac{1}{4} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{4} & 1 & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \end{bmatrix}, \quad A^{10} \approx \begin{bmatrix} 0.173 & 0.172 & 0.172 & 0.150 \\ 0.344 & 0.345 & 0.301 & 0.344 \\ 0.247 & 0.247 & 0.270 & 0.236 \\ 0.236 & 0.236 & 0.258 & 0.270 \end{bmatrix},$$

$$A^{20} \approx \begin{bmatrix} 0.16697 & 0.16697 & 0.16650 & 0.16623 \\ 0.33347 & 0.33347 & 0.33246 & 0.33393 \\ 0.25008 & 0.25008 & 0.25035 & 0.24948 \\ 0.24948 & 0.24948 & 0.25068 & 0.25035 \end{bmatrix}.$$

These powers A^m will be transition matrices as well; see Exercise 68.

In Exercises 69 through 72, you will have a chance to explore the significance of the entries of these matrices A^m, in terms of our mini-Web and its graph.

As we take a closer look at the matrix A^{20}, our attention may be drawn to the fact that the four column vectors are all close to the vector

$$\begin{bmatrix} 1/6 \\ 1/3 \\ 1/4 \\ 1/4 \end{bmatrix},$$

which happens to be the equilibrium distribution $\vec{x}_{equ}$ for the matrix A, as c
in Example 2.1.9. We might conjecture that the *limit* of the column vecto
is $\vec{x}_{equ}$ as m goes to infinity.

Before we address this issue, let's introduce some terminology.

Definition 2.3.10 **Regular transition matrices**

A transition matrix is said to be *positive* if all its entries are positive (i
that all the entries are greater than 0).

A transition matrix is said to be *regular* (or *eventually positive*) if th
A^m is positive for some positive integer m.

For example, the transition matrix $\begin{bmatrix} 0.4 & 0.7 \\ 0.6 & 0.3 \end{bmatrix}$ is positive (and ther
regular; let $m = 1$ in Definition 2.3.10). The transition matrix $A = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$
to be positive, but it is regular since $A^2 = \begin{bmatrix} 1/2 & 1/4 \\ 1/2 & 3/4 \end{bmatrix}$.

The matrix A in Example 4 fails to be positive, but it is regular si
positive. The reflection matrix $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ fails to be regular since A^m
for even m and $A^m = A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ for odd m.

Now we can address the conjecture we made at the end of Example

Theorem 2.3.11 **Equilibria for regular transition matrices**

Let A be a regular transition matrix of size $n \times n$.

a. There exists exactly one distribution vector $\vec{x}$ in $\mathbb{R}^n$ such that
This is called the *equilibrium distribution* for A, denoted $\vec{x}_{e}$
components of $\vec{x}_{equ}$ are positive.

b. If $\vec{x}$ is any distribution vector in $\mathbb{R}^n$, then $\lim_{m\to\infty}(A^m\vec{x}) = \vec{x}_{equ}$.

c. $\lim_{m\to\infty} A^m = \begin{bmatrix} | & & | \\ \vec{x}_{equ} & \cdots & \vec{x}_{equ} \\ | & & | \end{bmatrix}$, which is the matrix whose colu
$\vec{x}_{equ}$.[9]

Part (b) states that in the long run the system will approach the e
distribution $\vec{x}_{equ}$, *regardless of the initial distribution*; we say that
globally stable equilibrium distribution.

We will outline a proof of parts (a) and (b) in Chapter 7. I
(c) are easily seen to be equivalent. If we assume that part (b)
$\lim_{m\to\infty}$ (jth column of A^m) $= \lim_{m\to\infty}(A^m\vec{e}_j) = \vec{x}_{equ}$ since $\vec{e}_j$ is a distribu
Exercise 73, you are asked to derive part (b) from part (c).

[9]This limit is defined entry by entry. We claim that any entry of A^m converges to the
entry of the matrix $\begin{bmatrix} | & & | \\ \vec{x}_{equ} & \cdots & \vec{x}_{equ} \\ | & & | \end{bmatrix}$ as m goes to infinity.

EXERCISES 2.3

GOAL *Compute matrix products column by column and entry by entry. Interpret matrix multiplication in terms of the underlying linear transformations. Use the rules of matrix algebra. Multiply block matrices.*

If possible, compute the matrix products in Exercises 1 through 13, using paper and pencil.

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

2. $\begin{bmatrix} 1 & -1 \\ 0 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 1 & 0 \end{bmatrix}$

3. $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

4. $\begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} 7 & 5 \\ 3 & 1 \end{bmatrix}$

5. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

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

7. $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 1 & -1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 2 & 1 & 3 \end{bmatrix}$

8. $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

9. $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} -6 & 8 \\ 3 & -4 \end{bmatrix}$

10. $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$

11. $\begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$

12. $\begin{bmatrix} 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 1 \end{bmatrix}$

13. $\begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

14. For the matrices

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad E = \begin{bmatrix} 5 \end{bmatrix},$$

determine which of the 25 matrix products AA, AB, $AC, \ldots, ED$, EE are defined, and compute those that are defined.

Use the given partitions to compute the products in Exercises 15 and 16. Check your work by computing the same products without using a partition. Show all your work.

15. $\left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 1 & 3 & 4 \end{array}\right] \left[\begin{array}{c|c} 1 & 0 \\ 2 & 0 \\ \hline 3 & 4 \end{array}\right]$

16. $\left[\begin{array}{cc|cc} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \left[\begin{array}{cc|cc} 1 & 2 & 2 & 3 \\ 3 & 4 & 4 & 5 \\ \hline 0 & 0 & 1 & 2 \\ 0 & 0 & 3 & 4 \end{array}\right]$

In the Exercises 17 through 26, find all matrices that commute with the given matrix A.

17. $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$

18. $A = \begin{bmatrix} 2 & 3 \\ -3 & 2 \end{bmatrix}$

19. $A = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}$

20. $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$

21. $A = \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$

22. $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

23. $A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$

24. $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

25. $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

26. $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}$

27. Prove the *distributive laws* for matrices:

$$A(C + D) = AC + AD$$

and

$$(A + B)C = AC + BC.$$

28. Consider an $n \times p$ matrix A, a $p \times m$ matrix B, and a scalar k. Show that

$$(kA)B = A(kB) = k(AB).$$

29. Consider the matrix

$$D_\alpha = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}.$$

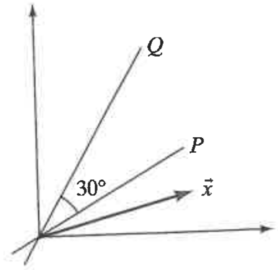
We know that the linear transformation $T(\vec{x}) = D_\alpha \vec{x}$ is a counterclockwise rotation through an angle α.

a. For two angles, α and β, consider the products $D_\alpha D_\beta$ and $D_\beta D_\alpha$. Arguing geometrically, describe the linear transformations $\vec{y} = D_\alpha D_\beta \vec{x}$ and $\vec{y} = D_\beta D_\alpha \vec{x}$. Are the two transformations the same?

b. Now compute the products $D_\alpha D_\beta$ and $D_\beta D_\alpha$. Do the results make sense in terms of your answer in part (a)? Recall the trigonometric identities

$$\sin(\alpha \pm \beta) = \sin\alpha\cos\beta \pm \cos\alpha\sin\beta$$
$$\cos(\alpha \pm \beta) = \cos\alpha\cos\beta \mp \sin\alpha\sin\beta.$$

30. Consider the lines P and Q in $\mathbb{R}^2$ in the accompanying figure. Consider the linear transformation $T(\vec{x}) = \text{ref}_Q(\text{ref}_P(\vec{x}))$; that is, we first reflect $\vec{x}$ about P and then we reflect the result about Q.



a. For the vector $\vec{x}$ given in the figure, sketch $T(\vec{x})$. What angle do the vectors $\vec{x}$ and $T(\vec{x})$ enclose? What is the relationship between the lengths of $\vec{x}$ and $T(\vec{x})$?

b. Use your answer in part (a) to describe the transformation T geometrically, as a reflection, rotation, shear, or projection.

c. Find the matrix of T.

d. Give a geometrical interpretation of the linear transformation $L(\vec{x}) = \text{ref}_P(\text{ref}_Q(\vec{x}))$, and find the matrix of L.

31. Consider two matrices A and B whose product AB is defined. Describe the ith row of the product AB in terms of the rows of A and the matrix B.

32. Find all 2×2 matrices X such that $AX = XA$ for all 2×2 matrices A.

For the matrices A in Exercises 33 through 42, compute $A^2 = AA$, $A^3 = AAA$, and A^4. Describe the pattern that emerges, and use this pattern to find A^{1001}. Interpret your answers geometrically, in terms of rotations, reflections, shears, and orthogonal projections.

33. $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ 34. $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ 35. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

36. $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ 37. $\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$ 38. $\frac{1}{2}\begin{bmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix}$

39. $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ 40. $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

41. $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ 42. $\frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

In Exercises 43 through 48, find a 2×2 matrix A with the given properties. Hint: It helps to think of geometrical examples.

43. $A \neq I_2$, $A^2 = I_2$ 44. $A^2 \neq I_2$, $A^4 = I_2$

45. $A^2 \neq I_2$, $A^3 = I_2$

46. $A^2 = A$, all entries of A are nonzero.

47. $A^3 = A$, all entries of A are nonzero.

48. $A^{10} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

In Exercises 49 through 54, consider the matrice

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

$D = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$, $E = \begin{bmatrix} 0.6 & 0.8 \\ 0.8 & -0.6 \end{bmatrix}$, $F = [$

$G = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, $H = \begin{bmatrix} 0.8 & -0.6 \\ 0.6 & 0.8 \end{bmatrix}$, $J = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

Compute the indicated products. Interpret th ucts geometrically, and draw composition diagr Example 2.

49. AF and FA 50. CG and GC

51. FJ and JF 52. JH and HJ

53. CD and DC 54. BE and EB.

In Exercises 55 through 64, find all matrices X t the given matrix equation.

55. $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} X = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

56. $X \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} = I_2$ 57. $\begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} X =$

58. $X \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ 59. $X \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} =$

60. $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} X = I_2$ 61. $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{bmatrix}$

62. $\begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 2 \end{bmatrix} X = I_3$ 63. $\begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} X$

64. $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} X = I_2$

65. Find all upper triangular 2×2 matrices X is the zero matrix.

66. Find all lower triangular 3×3 matrices X is the zero matrix.

67. a. If A is any 3×3 transition matrix tion 2.1.4), find the matrix product [1

b. For a fixed n, let $\vec{e}$ be the row $[1 \ 1 \ \ldots \ 1]$ (n 1's). Show that an $n \times n$ nonnegative entries is a transition ma only if) $\vec{e}A = \vec{e}$.

68. Show that if A and B are $n \times n$ transition AB will be a transition matrix as well. H cise 67b.

69. Consider the matrix A^2 in Example 4 of

a. The third component of the first colur What does this entry mean in practica in terms of surfers following links in

b. When is the ijth entry of A^2 equal to 0? Give your answer both in terms of paths of length 2 in the graph of the mini-Web and also in terms of surfers being able to get from page j to page i by following two consecutive links.

70. **a.** Compute A^3 for the matrix A in Example 2.3.4.

b. The fourth component of the first column of A^3 is 1/4. What does this entry mean in practical terms, that is, in terms of surfers following links in our mini-Web?

c. When is the ijth entry of A^3 equal to 0? Give your answer both in terms of paths in the graph of the mini-Web and also in terms of surfers being able to get from page j to page i by following consecutive links.

d. How many paths of length 3 are there in the graph of the mini-Web from page 1 to page 2? How many surfers are taking each of these paths, expressed as a proportion of the initial population of page 1?

71. For the mini-Web in Example 2.3.4, find pages i and j such that it is impossible to get from page j to page i by following exactly four consecutive links. What does the answer tell you about the entries of A^4?

72. For the mini-Web in Example 2.3.4, find the smallest positive integer m such that all the entries of A^m are positive; you may use technology. What does your answer tell you in terms of paths in the graph of the mini-Web and also in terms of surfers following consecutive links?

73. Use part (c) of Theorem 2.3.11 to prove part (b): If A is a regular transition matrix of size $n \times n$ with equilibrium distribution $\vec{x}_{equ}$, and if $\vec{x}$ is any distribution vector in $\mathbb{R}^n$, then $\lim_{m\to\infty}(A^m\vec{x}) = \vec{x}_{equ}$.

74. Suppose A is a transition matrix and B is a positive transition matrix (see Definition 2.3.10), where A and B are of the same size. Is AB necessarily a positive transition matrix? What about BA?

75. Prove the following: If A is a transition matrix and A^m is positive, then A^{m+1} is positive as well.

76. For the mini-Web graphed below, find the equilibrium distribution in the following way: Write the transition matrix A, test high powers of A to come up with a conjecture for the equilibrium distribution $\vec{x}_{equ}$, and then verify that $A\vec{x}_{equ} = \vec{x}_{equ}$. (This method, based on Theorem 2.3.11, is referred to as the *power method* for finding the equilibrium distribution of a regular transition matrix.) Also, find the page with the highest naïve PageRank. You may use technology.

$$\begin{array}{ccc} 1 & \rightleftarrows & 2 \\ & \nearrow & \downarrow \\ 3 & \leftarrow & 4 \end{array}$$

77. Consider the transition matrix

$$A = \begin{bmatrix} 0.4 & 0.2 & 0.7 \\ 0 & 0.6 & 0.1 \\ 0.6 & 0.2 & 0.2 \end{bmatrix}.$$

Verify that A is a regular transition matrix and then use the power method (see Exercise 76) to find the equilibrium distribution. You may use technology.

78. Let's revisit the mini-Web with the graph

$$\begin{array}{ccc} 1 & \rightleftarrows & 2 \\ \downarrow & \swarrow & \uparrow \\ 3 & \rightarrow & 4, \end{array}$$

but here we consider the surfing model with a "jumping rate" of 20%, as discussed in Exercise 2.1.53. The corresponding transition matrix is

$$B = \begin{bmatrix} 0.05 & 0.45 & 0.05 & 0.05 \\ 0.45 & 0.05 & 0.05 & 0.85 \\ 0.45 & 0.45 & 0.05 & 0.05 \\ 0.05 & 0.05 & 0.85 & 0.05 \end{bmatrix}.$$

This transition matrix is positive and therefore regular, so that Theorem 2.3.11 applies. Use the power method (see Exercise 76) to find the equilibrium distribution. You may use technology. Write the components of $\vec{x}_{equ}$ as rational numbers.

79. Give an example of a transition matrix A such that there exists more than one distribution vector $\vec{x}$ with $A\vec{x} = \vec{x}$.

80. Give an example of a transition matrix A such that $\lim_{m\to\infty} A^m$ fails to exist.

81. If $A\vec{v} = 5\vec{v}$, express $A^2\vec{v}$, $A^3\vec{v}$, and $A^m\vec{v}$ as scalar multiples of the vector $\vec{v}$.

82. In this exercise we will verify part (b) of Theorem 2.3.11 in the special case when A is the transition matrix $\begin{bmatrix} 0.4 & 0.3 \\ 0.6 & 0.7 \end{bmatrix}$ and $\vec{x}$ is the distribution vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. [We will not be using parts (a) and (c) of Theorem 2.3.11.] The general proof of Theorem 2.3.11 runs along similar lines, as we will see in Chapter 7.

a. Compute $A\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $A\begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Write $A\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ as a scalar multiple of the vector $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

b. Write the distribution vector $\vec{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ as a linear combination of the vectors $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

c. Use your answers in parts (a) and (b) to write $A\vec{x}$ as a linear combination of the vectors $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$. More generally, write $A^m\vec{x}$ as a linear combination of the vectors $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$, for any positive integer m. See Exercise 81.

d. In your equation in part (c), let m go to infinity to find $\lim_{m\to\infty} (A^m\vec{x})$. Verify that your answer is the equilibrium distribution for A.

83. If $A\vec{x} = \vec{x}$ for a regular transition matrix A and a distribution vector $\vec{x}$, show that all components of $\vec{x}$ must be positive. (Here you are proving the last claim rem 2.3.11a.)

84. Consider an $n \times m$ matrix A of rank n. Show exists an $m \times n$ matrix X such that $AX = I_n$. how many such matrices X are there?

85. Consider an $n \times n$ matrix A of rank n. How m matrices X are there such that $AX = I_n$?

2.4 The Inverse of a Linear Transformation

Let's first review the concept of an invertible function. As you read thes definitions, consider the examples in Figures 1 and 2, where X and Y are f

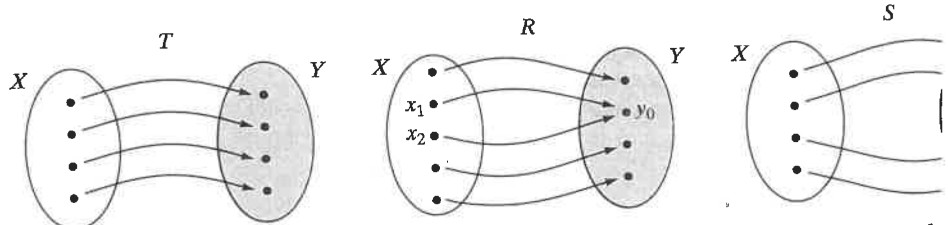


Figure 1 T is invertible. R is not invertible: The equation $R(x) = y_0$ has two solutions, x_1 and invertible: There is no x such that $S(x) = y_0$.

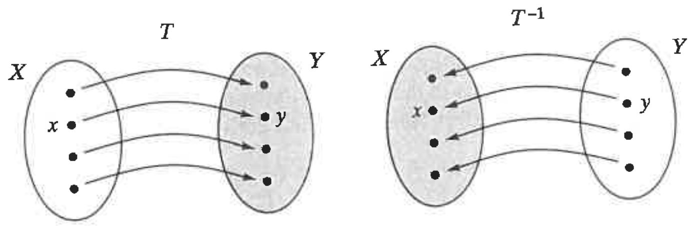


Figure 2 A function T and its inverse T^{-1}.

Definition 2.4.1 **Invertible Functions**

A function T from X to Y is called invertible if the equation $T(x)$ unique solution x in X for each y in Y.

In this case, the inverse T^{-1} from Y to X is defined by

$$T^{-1}(y) = \big(\text{the unique } x \text{ in } X \text{ such that } T(x) = y\big).$$

To put it differently, the equation

$$x = T^{-1}(y) \qquad \text{means that} \qquad y = T(x).$$

Note that

$$T^{-1}\big(T(x)\big) = x \qquad \text{and} \qquad T\big(T^{-1}(y)\big) = y$$

for all x in X and for all y in Y.

Conversely, if L is a function from Y to X such that

$$L\big(T(x)\big) = x \qquad \text{and} \qquad T\big(L(y)\big) = y$$

for all x in X and for all y in Y, then T is invertible and $T^{-1} = L$.

If a function T is invertible, then so is T^{-1} and $(T^{-1})^{-1} = T$.