Matrices

The Algebra of Multiple Values

Let's look at two problems: one with a single value and another very similar one with two values.

Single Value	Two Values
There is an island, with families living on it.	There is an island, with two kinds of families living on it: families that had their kids early, and families that had their kids late.
Each family has on average 4 members.	Families that started early have on average 2 adults and 3 children; families that started late have on average 2 adults and 2 children.
(so if *f* is the number of families, the number of family members is $m = 4f$)	(so if *e* is the number of families that started early and *l* is the number that started late, the numbers of adults and children are $a = 2e + 2l$) $c = 3e + 2l$
Each person eats 3000 calories of food per day.	Adults eat 1000 calories of protein and 2000 calories of carbohydrates daily; children eat 800 calories of protein and 1800 calories of carbohydrates daily.
(so total calories consumed is $d = 3000m$)	(so total calories consumed is $p = 1000a + 800c$) $h = 2000a + 1800c$
The island produces 528000 calories daily. How many families can it support with nothing left over?	The island produces 174400 calories of protein and 370400 calories of carbohydrates daily. How many families can it support with nothing left over?
$528000 = 3000 \cdot 4 \cdot f$, solve for *f*	?! *shoot me now*

Single vs. Multiple values

Once we get over the fact that it's a word problem, the left-hand problem is easy because we have a language and a notation that allow us to express the whole problem in that last line, which we can then solve. With the right-hand problem we can't do that. We have lots of formulas spread over the page. Even if we combine them all together and eventually solve them, we don't have the same level of understanding that we get from having it all in one formula.

More than that: on the left-hand problem, with everything in one formula, I can ask a question like

> Suppose there are 50 families on the island - how many calories can each person consume?

and you would write $528000 = c \cdot 4 \cdot 50$ and solve for *c*.

But if I try to do the same thing with the right-hand problem:

> Suppose there are 30 early families and 20 late families on the island - how many calories of protein and carbohydrate can each adult and child consume?

you have a problem that will put you in tears (try to solve it if you dare).

What we really need is a *notation* that will let us write single formulas that handle many values, and *rules of algebra* that will let us work with those formulas. Let's develop those things.

Vectors - One Name For Multiple Values

A ***vector*** is a list of numbers. We write the numbers in a column between big brackets: $\begin{bmatrix}1\\4\\2\\3\end{bmatrix}$ or $\begin{bmatrix}2\\b\end{bmatrix}$, for example. We write the name of a vector in boldface lowercase, like so: $\mathbf{v}$. We will still have need for ordinary numbers, whose names we will continue to write in plain italics, like *k*.

The individual numbers making up a vector are called the vector's ***components***. We refer to them by subscripts on the name of the vector, so that if $\mathbf{v}=\begin{bmatrix}1\\3\\5\end{bmatrix}$, $v_1=1$ and $v_2=3$.

Labeled Vectors

The components of a vector don't have to measure the same kind of thing. In other words, they don't have to have the same units. Often, we will be writing the units over to the left of the vector, to remind us of what they are. The vectors appearing in the opening problem would then be written as $\begin{matrix}\mathit{adult}\\\mathit{child}\end{matrix}\begin{bmatrix}2\\2\end{bmatrix}$ and $\begin{matrix}\mathit{protein}\\\mathit{carbs}\end{matrix}\begin{bmatrix}174400\\370400\end{bmatrix}$. These ***labeled vectors*** are just 'training wheels' for learning vectors - you never see them in standard mathematical use.

Adding Vectors

Adding vectors applies when you have a question like,

> Town A contains 2000 people and 1000 pets; town B has 5000 people and 500 pets; how many people and pets do the two towns have together?

To ***add two vectors***, just add them component by component to produce a vector of the same length. $\begin{matrix}\mathit{people}\\\mathit{pet}\end{matrix}\begin{bmatrix}2000\\1000\end{bmatrix}+\begin{matrix}\mathit{people}\\\mathit{pet}\end{matrix}\begin{bmatrix}5000\\500\end{bmatrix}=\begin{matrix}\mathit{people}\\\mathit{pet}\end{matrix}\begin{bmatrix}7000\\1500\end{bmatrix}$. It's really as simple as it looks.

Other examples: if $\mathbf{f} = \begin{matrix} early \\ late \end{matrix} \begin{bmatrix} e \\ l \end{bmatrix}$, $\mathbf{g} = \begin{matrix} early \\ late \end{matrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ then $\mathbf{f} + \mathbf{g} = \begin{matrix} early \\ late \end{matrix} \begin{bmatrix} e+2 \\ l+3 \end{bmatrix}$. If $\mathbf{u} = \begin{bmatrix} 2 \\ -1 \\ 6 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 3 \\ 4 \\ -1 \end{bmatrix}$ then $\mathbf{u} + \mathbf{v} = \begin{bmatrix} 5 \\ 3 \\ 5 \end{bmatrix}$.

With vectors you get two new ways of making a mistake in addition. First, you can try to add vectors that don't have the same number of components; this is meaningless since you have no way to compute the components of the result. $\begin{bmatrix} 2 \\ 5 \end{bmatrix} + \begin{bmatrix} 0 \\ 2 \\ 4 \end{bmatrix}$, $\begin{bmatrix} 2 \\ 4 \end{bmatrix} + 5$, and $\begin{bmatrix} 2 \end{bmatrix} + \begin{bmatrix} 4 \\ 2 \\ 7 \end{bmatrix}$ are all invalid.

Second, you have a whole new way of making the mistake of adding apples and oranges. You need to make sure that the corresponding components of the vectors you are adding have the same units. While it is true that $\begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 7 \\ 10 \end{bmatrix}$, if the labels are $\begin{matrix} apple \\ orange \end{matrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{matrix} orange \\ apple \end{matrix} \begin{bmatrix} 5 \\ 6 \end{bmatrix}$, what could the labels for the result be?

Vector Equality

Two vectors are equal if all corresponding components are equal. Put another way, **a=b** if **a-b=0** where **0** is the ***zero vector***, the vector all of whose components are 0.

What Is Multiplication?

You have been doing multiplication so long that you don't think about what it means any more. Well, you have to think again, because it turns out that you have been using multiplication to do two very different things, and those things will require different actions when you are operating on vectors.

Multiplication: Making Copies (*Scaling*)

The first meaning of multiplication is seen in a problem like:

> Each person has 3 fish. How many fish do 4 people have?

Here, you are just taking one person – 3 fish – and making 4 copies of them. It's the same as if you had just added repeatedly, 3 fish+3 fish+3 fish+3 fish.

A similar problem with multiple values would be:

> Each person has 3 apples and 4 oranges. How many apples and oranges do 4 people have?

Here you are making 4 copies of the vector $\begin{matrix} apple \\ orange \end{matrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix}$. **The key point is that the number of copies is a number, not a vector**. You just want to make 4 copies of each component. To do that, you multiply each component by the number: $4 \begin{matrix} apple \\ orange \end{matrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{matrix} apple \\ orange \end{matrix} \begin{bmatrix} 4 \cdot 3 \\ 4 \cdot 4 \end{bmatrix} = \begin{matrix} apple \\ orange \end{matrix} \begin{bmatrix} 12 \\ 16 \end{bmatrix}$.

This is the first kind of multiplication with vectors: multiplying a number times a vector. The number has no units, so the resulting vector has the same units as the original vector. You just multiply each component of the vector by the number. This is called ***scalar multiplication*** or simply ***scaling***.

Examples: if $\mathbf{v} = \begin{matrix} protein \\ carbs \end{matrix} \begin{bmatrix} 174400 \\ 370400 \end{bmatrix}$ then $2\mathbf{v} = 2 \begin{matrix} protein \\ carbs \end{matrix} \begin{bmatrix} 174400 \\ 370400 \end{bmatrix} = \begin{matrix} protein \\ carbs \end{matrix} \begin{bmatrix} 348800 \\ 740800 \end{bmatrix}$. If $\mathbf{f} = \begin{bmatrix} e \\ l \end{bmatrix}$ then $k\mathbf{f} = \begin{bmatrix} ke \\ kl \end{bmatrix}$.

Multiplication: Transformations

The other face of multiplication is seen in a problem like the following:

> You are traveling at 60 miles/hour. How far do you go in 3 hours?

This too is a multiplication problem, but you are not doing the same thing as you were with the people with fish. You are not making 60 copies of 3 hours, or 3 copies of 60 miles/hour. Rather, **you are transforming miles to hours. You are replacing each hour with 60 miles**. There are 3 hours, so when you change each hour to 60 miles you end up with 3×60=180 miles total. This kind of multiplication is called ***transformation.***

You might be thinking, “So what? In both cases I multiply; I can leave the why to the philosophers.” Alas, not so.

When you enter the land of multiple values you have to distinguish the two kinds of multiplication. Scaling is used when you are simply making copies, i. e. you are multiplying by a pure number without units, or when every component of the vector is modified in the same way. An example of the latter would be:

> Car A is driving at 50 miles/hour, and car B at 60 miles/hour. How far do they go in 3 hours?
>
> Solution: $3\ hours \cdot \begin{matrix} A\ mi/hr \\ B\ mi/hr \end{matrix} \begin{bmatrix} 50 \\ 60 \end{bmatrix} = \begin{matrix} A\ mi \\ B\ mi \end{matrix} \begin{bmatrix} 150 \\ 180 \end{bmatrix}$

Transformation is much more interesting, as in our example problem:

> Families that started early have on average 2 adults and 3 children; families that started late have on average 2 adults and 2 children.

Here **each of the multiple values** (early families, late families) **gets turned into a combination of the result values** (adults, children). There's not just one number to multiply by, like when we converted hours to miles; there are four.

We will need more than our simple scalar multiplication to perform transformation, and we need one new concept.

Linear Combinations

Your grade in this class is computed as 85% from tests, 10% from homework, and 5% from participation in class. If your test average is 94, your homework average is 90 (you got 100 on all you turned in, but you took one zero out of the 10 assignments), and you keep to yourself, earning an 80 on participation, what is your grade?

You got $0.85(94) + 0.1(90) + 0.05(80) = 92.9$ (B), right? Congratulations! You computed a ***linear combination***.

We will mostly be concerned with linear combinations of vectors. Suppose David Diligent got 92 on the tests, but turned in all the homework and chipped in on every question, earning 100 in both categories. How would the teacher compute your averages?

Rather than calculating each average separately, the teacher could do it all at once, with a ***linear combination of vectors***: $0.85\begin{bmatrix}95\\92\end{bmatrix} + 0.1\begin{bmatrix}90\\100\end{bmatrix} + 0.05\begin{bmatrix}80\\100\end{bmatrix} = \begin{matrix}\textit{you}\\\textit{David}\end{matrix}\begin{bmatrix}92.9\\93.2\end{bmatrix}$. [Oooh, too bad about that missed homework!]

A linear combination of vectors takes a set of vectors (all with the same number of components) and a set of numbers, multiplies the first vector by the first number, the second vector by the second number, the third… and so on, and then adds all the products together. The final result is a vector with the same shape as one of the original vectors. The numbers are called the ***weights***.

Writing this in formal mathematical notation, if you have a set of vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ and a set of weights $w_1, w_2, \ldots, w_n$, the linear combination is a vector given by $w_1\mathbf{u}_1 + w_2\mathbf{u}_2 + \ldots + w_n\mathbf{u}_n$.

Transforming Vectors - The Matrix Concept

Back to our problem

> Families that started early have on average 2 adults and 3 children;
> families that started late have on average 2 adults and 2 children.

When we wrote this out we got $\begin{matrix}a = 2e + 2l\\c = 2e + 3l\end{matrix}$. What we're looking for is a way to express that in a single formula. Right away we see that what comes out of the two equations is a vector $\begin{bmatrix}a\\c\end{bmatrix}$ of adults and children. And hidden inside the right-hand sides is the vector $\begin{bmatrix}e\\l\end{bmatrix}$ of early starters and late starters, and… all that other stuff. We have to figure out how to organize that other stuff.

Take a deep breath. This next section is the key to this chapter.

We see that each early family contributes 2 adults and 3 children. In other words, we have a vector $\begin{matrix} adult \\ child \end{matrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ and each early family is going to add that many people to the population. So we imagine replacing each early family by the vector $\begin{matrix} adult \\ child \end{matrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ - that's the transformation - and then we add up all those vectors, one copy for each of the *e* early families, which is just the scaling $e \begin{matrix} adult \\ child \end{matrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

Likewise, each late family is going to contribute $\begin{matrix} adult \\ child \end{matrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ to the total, making a total contribution of $l \begin{matrix} adult \\ child \end{matrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix}$. The final result, the vector $\begin{bmatrix} a \\ c \end{bmatrix}$, will be the total of those:

$$\begin{matrix} adult \\ child \end{matrix} \begin{bmatrix} a \\ c \end{bmatrix} = e \begin{matrix} adult \\ child \end{matrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} + l \begin{matrix} adult \\ child \end{matrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix}.$$

It is of surpassing importance that you understand that what we just did was calculate a linear combination of $\begin{matrix} adult \\ child \end{matrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ **and** $\begin{matrix} adult \\ child \end{matrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix}$**, using the vector components *e* and *l* as weights.**

Matrix Times Vector

We just performed a transformation of a vector $\begin{bmatrix} e \\ l \end{bmatrix}$ by making a linear combination of the two vectors $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 2 \end{bmatrix}$. This is such an important operation that we need a notation for it. First, we collect the vectors as the columns of a table, called a ***matrix***, written with one set of brackets around the entire matrix: $\begin{bmatrix} 2 & 2 \\ 3 & 2 \end{bmatrix}$. Then we define our desired linear combination to be the matrix $\begin{bmatrix} 2 & 2 \\ 3 & 2 \end{bmatrix}$ times the vector $\begin{bmatrix} e \\ l \end{bmatrix}$: $\begin{bmatrix} 2 & 2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} e \\ l \end{bmatrix} \equiv e \begin{bmatrix} 2 \\ 3 \end{bmatrix} + f \begin{bmatrix} 2 \\ 2 \end{bmatrix}$. **This is the definition of matrix times vector.**

A ***matrix*** is a rectangular table of numbers whose columns are treated as vectors.

Matrix times vector means "first number of the vector times the first column of the matrix, plus second number of the vector times the second column of the matrix plus… and so on until you are through".

The columns of the matrix are ordered to match the rows of the vector.

The result of each column-times-component multiplication must have the same units. But since the components of the vector may have different units, the columns of the matrix may need to have different units to make the column-times-component units come out right. When we label a matrix, we attach a label to each row indicating the units that will come out of the matrix-vector product, and we attach a label to each column indicating the units that the matrix expects the vector to have in the corresponding component. In this example, the matrix is $\begin{matrix} & \begin{matrix} early & late \end{matrix} \\ \begin{matrix} adult \\ child \end{matrix} & \begin{bmatrix} 2 & 2 \\ 3 & 2 \end{bmatrix} \end{matrix}$.

If there are 10 early families and 4 late families we calculate the total number of people as

$$\begin{matrix} & \begin{matrix} early & late \end{matrix} \\ \begin{matrix} adult \\ child \end{matrix} & \begin{bmatrix} 2 & 2 \\ 3 & 2 \end{bmatrix} \end{matrix} \bullet \begin{matrix} early \\ late \end{matrix} \begin{bmatrix} 10 \\ 4 \end{bmatrix} = 10 \begin{bmatrix} 2 \\ 3 \end{bmatrix} + 4 \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 20 \\ 30 \end{bmatrix} + \begin{bmatrix} 8 \\ 8 \end{bmatrix} = \begin{matrix} adult \\ child \end{matrix} \begin{bmatrix} 28 \\ 38 \end{bmatrix}$$

. If we tried to calculate

$$\begin{matrix} & \begin{matrix} early & late \end{matrix} \\ \begin{matrix} adult \\ child \end{matrix} & \begin{bmatrix} 2 & 2 \\ 3 & 2 \end{bmatrix} \end{matrix} \bullet \begin{matrix} late \\ early \end{matrix} \begin{bmatrix} 4 \\ 10 \end{bmatrix}$$

we would be making an error, because the row labels on the vector don't match the column labels on the matrix.

We write the name of a matrix in boldface uppercase, like so: **A**.

Pictorial View of Matrix-Vector Multiplication

To remember the pattern of matrix-vector multiplication, be sure to remember that everything is organized in columns – like columns of soldiers, perhaps. The soldiers in the vector march out, each armed with a bullhorn, and they march across the top of the matrix so that each one ends up at the head of one column of the matrix. Then each vector-soldier shouts his number to the matrix-column he has marched to, and the matrix-column is multiplied by that number. Finally, all the multiplied matrix-columns are added together to produce the result, which is a single

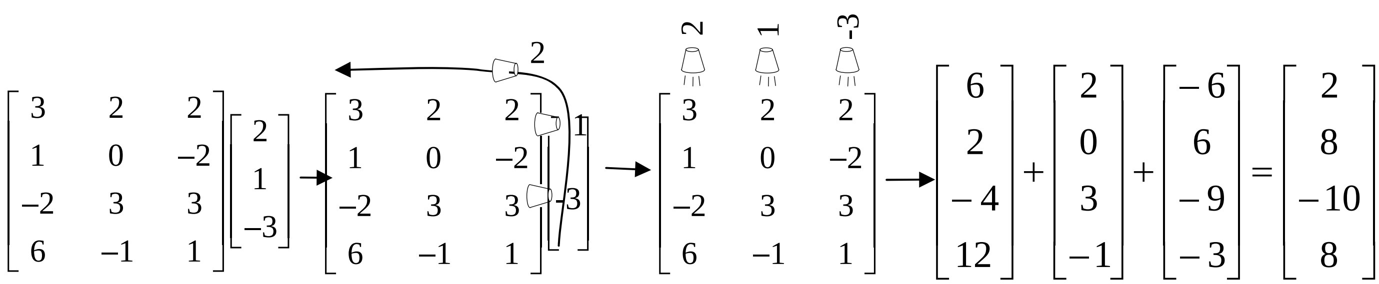


Figure 1. The sequence of matrix-vector multiplication

column.

Matrix Terms and Notation

When we describe the ***shape*** of a matrix (aka its ***dimensions***), that is, how many rows and columns it has, we give the number of rows first. We say that a matrix with 3 rows and 2 columns is a 3×2 matrix (read as 'three by two').

A matrix has two levels of organization: it is a set of column vectors, and each column has components. If we just want to focus on the column vectors we write the matrix as $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$ which shows that the matrix is made up of n column vectors, each with the same letter **a** and distinguished by subscripts. If we want to show the individual numbers of the matrix we write the $m \times n$ matrix as $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$ where each number has two subscripts showing the row and column it occupies. These individual numbers are called the ***components of the matrix***.

In a labeled matrix the column labels indicate what units are expected from the corresponding component of the vector, and the row labels indicate the units of the result of the multiplication.

Matrix Notation For Multi-Valued Formulas

Matrices give us the multi-valued notation we have been looking for! If **p** is the population vector (adults and children), and **f** is the families vector (early and late), and **S** is the family-size matrix (telling how many of each kind of person are in each kind of family), we can write one formula:

$$\mathbf{p} = \mathbf{S}\mathbf{f}$$

that tells how these things are related. We have a notation that is just as compact as what we used for single variables.

Just as with single-variable notation, we may replace the names with numbers if we have them. So we could write the equation above as $\mathbf{p} = \begin{bmatrix} 2 & 2 \\ 3 & 2 \end{bmatrix}\mathbf{f}$, and if someone tells us that the families vector $\mathbf{f} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$, we could substitute that in to get $\mathbf{p} = \begin{bmatrix} 2 & 2 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 14 \\ 19 \end{bmatrix}$.

Observations on Matrix-Vector Multiplication

We have discussed matrix-vector multiplication **Sf**. Don't think that this is the same as **fS**! We have not defined what, if anything, **fS** would mean.

Since the components of the column vector are going to be used as weights for the columns of the matrix, it follows that **the matrix must have the same number of columns as the vector has components**. If this is not the case, the multiplication is illegal and meaningless.

The result of multiplying a matrix times a column vector is always a column vector.

If you have labels on your matrix and vector, **the labels at the top of the matrix must match the labels on the vector, and the labels on the left of the matrix will be the labels on the result vector**.

The **columns of the matrix do not have to have the same length as the vector**. For example, if we know that early families tend to have 2 pets and late families have only 1, we can calculate the total number of people and pets as

$$\begin{array}{c} \\ adult \\ child \\ pet \end{array} \begin{array}{c} early\ \ late \\ \begin{bmatrix} 2 & 2 \\ 3 & 2 \\ 2 & 1 \end{bmatrix} \end{array} \begin{array}{c} early \\ late \end{array} \begin{bmatrix} 10 \\ 4 \end{bmatrix} = \begin{array}{c} adult \\ child \\ pet \end{array} \begin{bmatrix} 28 \\ 38 \\ 24 \end{bmatrix}.$$

The units inside the matrix do not have to be all the same, but they have to match with the units of the vector so that each component of the linear combination has the same units, which will allow those vectors to be added. In that last example the units of the matrix would have been

$$\begin{bmatrix} adult/early & adult/late \\ child/early & child/late \\ pet/early & pet/late \end{bmatrix}.$$

We can add as many new **rows** to the matrix as we like, but we cannot add any new **columns** without invalidating the multiplication. You may want to take a moment to look at that example and convince yourself that this makes sense, remembering what we said earlier about what it means to multiply a matrix by a vector.

Interpretation of Matrix-Vector Multiplication

Multiplying a vector by a matrix produces a ***transformation*** of the vector. It represents the vector in a different form.

Let's continue with the example:

$$\begin{array}{c} \\ adult \\ child \\ pet \end{array} \begin{array}{c} early\ \ late \\ \begin{bmatrix} 2 & 2 \\ 3 & 2 \\ 2 & 1 \end{bmatrix} \end{array} \begin{array}{c} early \\ late \end{array} \begin{bmatrix} 10 \\ 4 \end{bmatrix} = \begin{array}{c} adult \\ child \\ pet \end{array} \begin{bmatrix} 28 \\ 38 \\ 24 \end{bmatrix}$$

What does the vector $\begin{array}{c} early \\ late \end{array} \begin{bmatrix} 10 \\ 4 \end{bmatrix}$ represent? At a very basic level, it represents how big our island is. (Forgot we were talking about an island, didn't you?)

Now maybe there are many different islands of different sizes, but we're going to assume that, on small islands or big, it is still true that early families average 2 adults, 3 children, and 2 pets, and so on. This information is encoded in our matrix.

So the matrix allows you to take the question "how big is the island?" in its original form (number of early families, number of late families) and transform it into "how big is the island?" in a different form (adults, children, and pets). The matrix represents a conversion from one form of this information to another, but it really is the same information.

This is important—and subtle!—so let's try another example. Suppose an average adult book contains 250 pages, 12 illustrations, and weighs 2 pounds. An average children's book contains 30 pages, 30 illustrations (one on every page), and weighs 1 pound. I made all that up, of course—but if we pretend it's true, it offers a way of converting the size of my living-room book collection from one form (number of adult books, number of children's books) to a different form (number of pages, illustrations, and pounds).

So here's a good opportunity to test yourself.

1. If my living room has 250 adult books and 100 children's books, what is the size of my book collection measured in pages, illustrations, and pounds?
2. How can you express, and answer, my first question using matrices and vectors? (Make sure to use good labels!)
3. What would change, in your answer to #2, if we were talking about the Wake County Public Library instead of my living room? What would not change?

In the island example, we started with a vector representing families (early and late). Multiplying by the matrix ***transformed*** that into family members (adult, child, and pet).

Our basic understanding of matrices should be: **a matrix represents a transformation**.

The reasons why $\begin{bmatrix} a \\ c \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} e \\ f \end{bmatrix}$ is better than $\begin{matrix} a = 2e + 2l \\ c = 2e + 3l \end{matrix}$ are:

1. The matrix form pulls the transformation out into a unit, the matrix $\begin{bmatrix} 2 & 2 \\ 3 & 2 \end{bmatrix}$. In the multi-equation form the transformation and the values it operates on are intertwined.
2. With the values separated from the transformation, we can give a single name to each, and we can write the formula as $\mathbf{p} = \begin{bmatrix} 2 & 2 \\ 3 & 2 \end{bmatrix} \mathbf{f}$ or **p=Sf**.

Adding and Scaling Matrices

Scaling a matrix is the same as scaling each of its vectors:

$k\mathbf{A} = k\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} = \begin{bmatrix} k\mathbf{a}_1 & k\mathbf{a}_2 & \cdots & k\mathbf{a}_n \end{bmatrix}$. This is just saying that doing a transformation and then multiplying everything by *k* is the same as just doing a transformation that gives you *k* times as much of everything.

The statement **A=B+C** means that the transformation **A** does the same thing as doing the two transformations **B** and **C** and adding the results. If you consider the details you will see that this means that each vector of **A** must be equal to the sums of the corresponding vectors of **B** and **C**. Adding two matrices is the same as adding the corresponding vectors. It follows that two matrices must have the same shape if they are to be added.

The upshot of those two rules is that matrix scaling just multiplies each component by the number, and matrix addition just adds the corresponding components.

Matrices are equal if they have the same shape and all corresponding components are equal.

Multiplying Matrix Times Matrix

Referring again to our original problem, we found that we could calculate the number of persons from the number of families using the formula **p=Sf**.

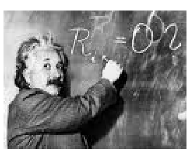
By a similar process of reasoning, we can see that the vector **c** of calories consumed (protein and carbohydrate) can be calculated from the numbers of persons **p** by multiplying by the consumption matrix $\mathbf{M} = \begin{matrix} & \textit{adult} & \textit{child} \\ \textit{protein} & 2 & 2 \\ \textit{carbs} & 3 & 2 \end{matrix}$, i.e. $\mathbf{M} = \begin{bmatrix} 2 & 2 \\ 3 & 2 \end{bmatrix}$ using the formula **c=Mp**.

Now, if we are going to do algebra with our vectors and matrices we need to be able to substitute freely; so substituting for **p** we get the overall formula

c=MSf

and we are now faced with the question of what **MSf** means. It surely can mean **M(Sf)**, in other words, multiply **Sf** first to get a vector, and then multiply by **M**; but it would be great if it could also mean (**MS**)**f**, which would mean that we would multiply **M** times **S** to get something - it would have to be a matrix - that would give the same result when multiplied by **f**.

If we can define **MS** that way, we will have a definition of multiplying matrix times matrix that will satisfy the associative property **M(Sf)=(MS)f**.



Definition of Matrix-Matrix Multiplication

Let's do it. We are going to find out how to define the matrix product **D=MS** so that **Df=M(Sf)**. Remembering that the matrices are made up of column vectors and the vector **f** is made up of numbers, we can write **Df=MSf** as

$$\begin{bmatrix} \mathbf{d}_1 & \mathbf{d}_2 & \cdots & \mathbf{d}_n \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix} = \mathbf{M} \begin{bmatrix} \mathbf{s}_1 & \mathbf{s}_2 & \cdots & \mathbf{s}_n \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix}$$

.Multiplying out the matrix-vector products, we get

$$f_1\mathbf{d}_1 + f_2\mathbf{d}_2 + ... + f_n\mathbf{d}_n = \mathbf{M}\left(f_1\mathbf{s}_1 + f_2\mathbf{s}_2 + ... + f_n\mathbf{s}_n\right).$$

But look: if this is going to be true for any combination of fs, it must be true when f_1=1 and all the other fs are 0; that leaves us with $\mathbf{d}_1 = \mathbf{M}f_1\mathbf{s}_1 = \mathbf{M}\mathbf{s}_1$, and the same idea applies when we set each f_i in turn to 1, leaving the others 0; in other words, $\mathbf{d}_1 = \mathbf{M}\mathbf{s}_1, \mathbf{d}_2 = \mathbf{M}\mathbf{s}_2, \ldots, \mathbf{d}_n = \mathbf{M}\mathbf{s}_n$ and we can write our final answer as

$$\mathbf{D} = \mathbf{MS} \equiv \begin{bmatrix} \mathbf{M}\mathbf{s}_1 & \mathbf{M}\mathbf{s}_2 & \cdots & \mathbf{M}\mathbf{s}_n \end{bmatrix}.$$ ***This is the definition of the product of two matrices***.

To put that into words, ***the product of two matrices* AB *is a matrix each of whose columns comes from the matrix* A *multiplied by the corresponding column of* B.**

Example of Labeled Matrix-Matrix Multiplication

Using the numbers of our example problem, we have

$$\mathbf{D} = \mathbf{MS} = \begin{matrix} & \textit{adult} & \textit{child} \\ \textit{protein} & 1800 & 800 \\ \textit{carbs} & 2000 & 1800 \end{matrix} \quad \begin{matrix} & \textit{early} & \textit{late} \\ \textit{adult} & 2 & 2 \\ \textit{child} & 3 & 2 \end{matrix} = \left[\begin{bmatrix} 1800 & 800 \\ 2000 & 1800 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} \; \begin{bmatrix} 1800 & 800 \\ 2000 & 1800 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} \right]$$

$$= \begin{matrix} & \textit{early} & \textit{late} \\ \textit{protein} & 4400 & 3600 \\ \textit{carbs} & 9400 & 7600 \end{matrix}.$$ This is saying that an early family needs 4400 calories of protein and 9400 calories of carbohydrates daily, while a late family needs 3600 calories of protein and 7600 calories of carbohydrates. Yeah, that seems right.

Observations on Matrix-Matrix Multiplication

Because a matrix-matrix multiplication is a sequence of matrix-vector multiplications, the same rules apply for the shapes of the matrices: each column in the right-hand matrix must have the same number of components as there are columns in the left-hand matrix. In other words, **the number of columns in the left-hand matrix must equal the number of rows in the right-hand matrix**.

The matrix product will have the same number of rows as the left-hand matrix and the same number of columns as the right-hand matrix. Multiplying an $m \times n$ matrix by a $n \times p$ matrix produces an $m \times p$ matrix.

If you have labels on your matrices, **the column labels of the left-hand matrix must match the row labels of the right-hand matrix, and these labels will not show up in the result. The row labels of the result will come from the left-hand matrix, and the column labels from the right-hand matrix.**

AB≠BA! **Matrix multiplication is not commutative!** This is pretty clear from the observations above: since the number of columns in the left operand and the number of rows in the right operand have to match, but the other dimensions are unconstrained, it might not even be possible to change the order of multiplication. And even if you could, the labels might not match. And anyway, even if they did, you usually get a different result: $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}$ but $\begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 4 \\ 1 & 0 \end{bmatrix}$. You're going to have to live with that.

Since matrix multiplication is not commutative, **get in the habit of thinking of your matrix products in right-to-left order**. You should normally view the product from the point of view of the vector that is being transformed. The matrix product **ABCu** starts with the vector **u**. First

C is applied to produce **Cu**, then **B** is applied to produce **B**(**Cu**), then **A** is applied to produce **A**(**B**(**Cu**)). If you see **ABC**, think to yourself "that's the transformation that you get from applying **C** followed by **B** followed by **A**".

Interpretation of Matrix-Matrix Multiplication

When we learned matrix-vector multiplication we learned that **a matrix represents a transformation**.

The product of matrices represents a sequence of transformations.

Look back at our example **D=MS**. **S**, the family-size matrix, transforms number of families to numbers of persons. **M**, the consumption matrix, transforms numbers of persons to calories per day. Their product **D**, the daily consumption matrix, is a single matrix that does both of these things: it converts numbers of families directly to calories per day.

Using labeled matrices makes the sequence clear. The intermediate labels disappear, and the product matrix converts from the input labels of the rightmost matrix to the output labels of the leftmost matrix.

Square Matrices And Operators

A matrix represents a transformation. Sometimes the transformation is a complete change, like going from families to calories; but in many cases of interest the output of the transformation is measured in the same units as the input. Such a transformation, that changes the numbers without changing the units, is called an ***operator***. Examples of operators are:

- taxation: you have dollars; the transformation is applied; you have fewer dollars.
- putting a filter on a camera: light, made up of red/green/blue components, comes in; the transformation is applied; a different combination of red/green/blue goes out.
- stretching, scaling, and rotating an object: object coordinates (x,y values) come in; the transformation is applied; modified object coordinates go out.

We will look at object stretching, scaling, and rotation because they are easy to depict graphically.

Like any matrix, the matrix representing an operator must have the same number of columns as there are components in the vectors it multiplies; but it must have that many rows too, because the result of the multiplication must have the same number of components as the vector. So, the matrix for an operator must have the same number of rows as columns. Such a matrix is called a ***square matrix***. There are some important things that square matrices can do that other matrices can't, and we will be exploring them presently.

Transformations In The Plane