

Introduction to Matrices

QEA

Spring 2016

What is this about?

The main mathematical tool that you will learn and use in this module is linear algebra. Much of linear algebra involves operations with matrices and vectors. While most of the examples you will encounter in this module will be from image processing, the tools you learn in this module are useful in a wide range of applications, including statistics, signal processing and numerous other topics. In this building block, you will learn the basics of matrix algebra, which will be the major tool you will use in Module 2. We shall start by introducing matrix operations on vectors, followed by matrix operations on other matrices.

Resources to read and watch

There are lots of books about Linear Algebra and lots of useful videos on the web. Here are some specific recommendations:

- Introduction to Linear Algebra, by Strang
- Linear Algebra, by Lay
- Homebrew videos
 - [Matrices operating on vectors](#)
 - [Matrices operating on vectors \(example\)](#)
 - [Matrices operating on matrices](#)

Matrices act on vectors

A MATRIX LOOKS LIKE AN ARRAY, BUT IT OPERATES ON VECTORS (AND OTHER MATRICES, AS WE WILL SOON SEE): when we multiply a vector by a matrix, it transforms the vector into a new vector. Therefore, when we say "a matrix operates on a vector", we mean that the matrix multiplies the vector. Notation-wise, we use bold upper-case letters, e.g. **A**, to represent a matrix and bold lower-case letters to represent a vector, e.g. **v**.

Matrices and vectors come in different shapes and sizes. Like arrays, we talk about matrices as having rows and columns. A general

matrix \mathbf{A} has m rows and n columns, and we refer to this as an $m \times n$ matrix. Vectors are then examples of matrices: they either have a single row or a single column, and we often refer to them as being a *row vector* or a *column vector*.

Matrices can only multiply vectors of a certain size and produce vectors of a certain size: an $m \times n$ matrix can only operate on a column vector of size $n \times 1$, and will produce an output vector which is a column vector of size $m \times 1$. Likewise, matrices can only multiply other matrices of a certain size: an $m \times n$ matrix can only act on a matrix of size $n \times k$, and will produce an output matrix of size $m \times k$. These basic properties will become clearer when we look at an example.

Example

Consider the 3×2 matrix \mathbf{A} ,

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

and the input vector \mathbf{v}

$$\mathbf{v} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

The output vector \mathbf{w} is computed as follows

$$\mathbf{w} = \begin{bmatrix} 2 & 1 \\ 3 & -1 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} (2)(-2) + (1)(1) \\ (3)(-2) + (-1)(1) \\ (0)(-2) + (4)(1) \end{bmatrix} = \begin{bmatrix} -3 \\ -7 \\ 4 \end{bmatrix}$$

There are two main ways to think about this multiplication. The most common view is to treat each entry of the new vector as a dot product between a row of the matrix and the column vector. So, for example, the first entry in the output vector is the dot product of two vectors

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

The second approach is to view the output vector as a linear combination of the columns of the matrix. The entries in the original vector are used as multiplication weights on each column of the matrix, i.e.

$$(-2) \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix} + (1) \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix} = \begin{bmatrix} -3 \\ -7 \\ 4 \end{bmatrix}$$

We encourage you to use both approaches when you think about multiplication.

General Notation

As we mentioned earlier, $n \times m$ matrices can multiply $m \times 1$ vectors and produce $n \times 1$ vectors. Consider a generic $n \times m$ matrix \mathbf{A}

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \cdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

where the ij -th entry of this matrix, a_{ij} defined above, is the entry corresponding to the i -th row and j -th column. You can multiply an $m \times 1$ vector \mathbf{v} by this matrix. Define the vector \mathbf{v} as follows,

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix}$$

Now define another vector \mathbf{w} which is the product of \mathbf{A} and \mathbf{v} , i.e. $\mathbf{w} = \mathbf{A}\mathbf{v}$. If we define

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$

then the i -th entry of \mathbf{w} , is given by the following sum

$$w_i = a_{i1}v_1 + a_{i2}v_2 + \cdots + a_{im}v_m = \sum_{j=1}^m a_{ij}v_j$$

Matrix Operations in MATLAB

MATLAB has a wide variety of tools to deal with matrix operations. You should open MATLAB on your computers and work through the following examples. To define the matrix \mathbf{A} in MATLAB you can type the following command:

```
>> A = [2 1; 3 -1; 0 4]
```

Note that the semi-colon ends a row and begins a new row. To define the vector \mathbf{v} in MATLAB you can type the following command:

```
>> v = [-2; 1]
```

To multiply \mathbf{A} and the vector \mathbf{v} in MATLAB you can type the following command:

```
>> w = A*v
```

Matrix Operations in Mathematica

Mathematica has a wide variety of tools to deal with matrix operations. You should open *Mathematica* on your computers and work through the following examples in a notebook. To define the matrix **a** in *Mathematica* you can manually type the following command (using a lowercase **a** to distinguish from builtin commands):

```
a = {{2, 1}, {3, -1}, {0, 4}} [shift-enter to evaluate]
```

In *Mathematica*, matrices are simply lists of lists, so all the operations that work to generate and manipulate lists (Reverse, Sort, etc) also work on matrices, and all the operations on matrices (Transpose, etc) also work on nested lists.

To avoid typing so many curly braces, you could use the **Palettes » Basic Math Assistant » Basic commands » Matrix commands** menu to insert a matrix graphically, or hover over that button to see the keyboard shortcuts.

To define the vector **v** in *Mathematica* you can type the following command:

```
v = {-2, 1}
```

To multiply **A** and the vector **v** in *Mathematica* you can type the following command:

```
w = a.v [shift-enter]
```

where "." refers to matrix multiplication, and "*" would refer to elementwise multiplication.

Extracting sub-matrices in MATLAB

Let's say we want to extract the 2nd row of the matrix **A**. In MATLAB you can type the following command:

```
>> A(2, :)
```

Note that the colon in this case simply grabs all of the columns. Alternatively, you can type the following command:

```
>> A(2, 1:2)
```

which will grab the first and second column, or

```
>> A(2, 1:end)
```

which will grab all the columns from 1 to the last column. If you want to grab the first and third rows only, you can type the following command:

```
>> A([1 3],:)
```

or just

```
>> A(1:2:3,:)
```

Extracting sub-matrices in Mathematica

Let's say we want to extract the 2nd row of the matrix **a**. In *Mathematica* you can type the following command:

```
a[[2, All]]
```

Note that the output of this command is a vector, not a matrix. Alternatively, you can type the following command:

```
a[[2, 1;;2]]
```

which will grab the first and second column, or

```
a[[2, 1;;All]]
```

which will grab all the columns from 1 to the last column. If you want to grab the first and third rows only, you can type the following command:

```
a[{{1, 3}, All}]
```

or just

```
a[[1;;3;;2, All]]
```

To make things look prettier, you can use the `MatrixForm[...]` command or **Display as » Matrix** button in the suggestion bar.

Exercises

- Using the definitions for **A** and **v** from above, please solve the following using MATLAB or *Mathematica*. Do the answers match what you expect? (Not all of these may be defined!)

MATLAB:

(a) » `A*v`

(b) » `A(1:2,:)*v`

(c) » `A(:, 2:4)*v`

Mathematica:

(a) `a.v`

(b) `a[[1;;2, All]].v`

(c) `a[[All, 2;;4]].v`

2. Find a 2×2 matrix which when multiplying a vector \mathbf{v} results in the same vector \mathbf{v} .
3. Find a 3×3 matrix which when multiplying a vector \mathbf{v} keeps the first two entries of \mathbf{v} the same and replaces the third entry with a zero.
4. Find a 4×4 matrix which when multiplying a vector \mathbf{v} swaps the first and second entries of the vector.

Matrices as transformation operators

WHEN MATRICES OPERATE ON SPATIAL POSITION VECTORS, THE VECTOR WHICH RESULTS IS ANOTHER SPATIAL POSITION VECTOR. The original spatial position has been 'transformed' into another position. In particular, there are specific matrices which accomplish particular desired transformations and are used in many different disciplines. We will use a couple of these as examples here.

Points as vectors, vectors as points

There is a natural correspondence between points in Euclidean space and vectors. Consider the point $(1, 2, -1)$ in 3D. We can associate a position vector with this point, which is just the vector from the origin to this point, namely

$$\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

Likewise, we can think of every vector as defining a point, if we assume that the vector emanates from the origin. So, for example, the vector

$$\mathbf{v} = \begin{bmatrix} 3 \\ -2 \\ 0 \\ 1 \end{bmatrix}$$

could be identified with the point $(3, -2, 0, 1)$ in 4D. Often times we will mix and match these ideas and say things like: the vector (x, y, z) . What we really mean when we say this is: the point (x, y, z) can be treated as the position vector

$$\mathbf{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$



Figure 1: Transformations, not Transformers.

Identity Operation

Ok, this is the easy one. The matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (1)$$

when it acts on the position vector (x, y, z) , will reproduce the same position. Convince yourself this is true!

Scaling Operation

Another important and simple operation is to be able to take a vector (x, y, z) and scale (increase or decrease its length) it by an overall multiplicative factor while maintaining its direction.

5. Consider the vector $(2, 0, 0)$. Thinking about how the identity matrix acts on this vector, propose a 3×3 matrix which scales this vector by a factor of 3 to the vector $(6, 0, 0)$.
6. Now think about the vector $(0, 2, 0)$. What 3×3 matrix can you propose which scales this vector by a factor of 3.
7. And finally, find the 3×3 matrix which scales the vector $(0, 0, 2)$ by a factor of 3.
8. Putting it together! What 3×3 matrix will scale the vector $(2, 2, 2)$ by a factor of 3.
9. What if you want to scale the x component differently than the y component? Write down the 3×3 matrix which scales the x component by 3 and the y component by 5 and leaves the z component the same.

One thing you might notice in working with the scaling is that all these operations only involve elements of your transformation matrix which lie on the diagonal. If you think about what the diagonal elements represent, they map the i^{th} component of the input vector onto the i^{th} component of the output vector, so this makes sense that diagonal elements would be heavily associated with scaling. The off-diagonal elements, in contrast, represent the mapping of the i^{th} component of the input vector onto the j^{th} component of the output vector, which will change the direction of the vector!

Rotation Operation

To start our exploration of the effects of off-diagonal elements, let's consider some simple examples.

10. Operate on the vectors $(1,0,0)$, $(0,1,0)$ and $(0,0,1)$ with the following matrix:

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2)$$

What transformation does this matrix represent? (ie. can you capture in a single concept what this matrix does to any vector? Hint: look at the section heading).

11. Similarly, consider the operation of the following matrix on the vectors $(1,0,0)$, $(0,1,0)$ and $(0,0,1)$:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \quad (3)$$

What transformation does this matrix represent?

12. Lastly, what transformation does this matrix represent?:

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix} \quad (4)$$

So, hopefully by this point you have convinced yourself that the above matrixes transform vectors by rotating them through an angle of 90 degrees about some coordinate axis. This is accomplished by the presence of the 1's and -1's in the off-diagonal positions which replace the i^{th} component of the vector with the j^{th} and vice versa with a negative sign. What if we want to rotate the vector through an angle which is not 90 degrees? We want to move some proportion of the i^{th} component: how much depends on the angle! Consider the following matrix:

$$\mathbf{R} = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (5)$$

13. What is this matrix equal to for $\theta = 0$? What is it equal to for $\theta = 90$ degrees?
14. If $\theta = \pi/2$ what is the result of operating with this matrix on the vector $(1,0,0)$?
15. If, for any θ , we operate on the vector (a,b,c) , show that the length (norm) of the vector is unchanged by this operation! (Hint: remember that $\cos^2 \theta + \sin^2 \theta = 1$).
16. What transformation does this matrix represent? Can you, by analogy, construct the rotation matrixes for rotation by an angle θ around the other two coordinate axes?

Matrices acting on matrices

Matrices can also operate on matrices to produce other matrices. In general, when you multiply a matrix \mathbf{A} with another matrix \mathbf{B} , you need the matrix on the left side of the product to have the same number of columns as the number of rows in the matrix on the right side. In other words if \mathbf{A} is $n \times m$, and \mathbf{B} is $p \times q$, you need $m = p$ for the product $\mathbf{C} = \mathbf{AB}$ to be defined. The product results in a new matrix \mathbf{C} which is $n \times q$. The q columns of the product matrix \mathbf{C} are precisely the q vectors that would result from multiplying \mathbf{A} with the vectors formed by the columns of \mathbf{B} .

Example 1

Consider the matrix \mathbf{A} ,

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

and another matrix \mathbf{B}

$$\mathbf{B} = \begin{bmatrix} 1 & 5 \\ -2 & 3 \end{bmatrix}$$

The product of the two $\mathbf{C} = \mathbf{AB}$ is computed as follows

$$\mathbf{C} = \begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 1 & 5 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} (2)(1) + (1)(-2) & (2)(5) + (1)(3) \\ (3)(1) + (-1)(-2) & (3)(5) + (-1)(3) \end{bmatrix} = \begin{bmatrix} 0 & 13 \\ 5 & 12 \end{bmatrix}$$

Example 2

As a second example consider the matrices \mathbf{A} and \mathbf{B} defined below, and let the product $\mathbf{C} = \mathbf{AB}$.

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

$$\mathbf{B} = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} (1)(1) + (2)(2) & (1)(4) + (2)(3) \\ (3)(1) + (2)(2) & (3)(4) + (2)(3) \\ (4)(1) + (1)(2) & (4)(4) + (1)(3) \end{bmatrix} = \begin{bmatrix} 5 & 10 \\ 7 & 18 \\ 6 & 19 \end{bmatrix}$$

More generally, consider an $n \times m$ matrix \mathbf{A} , a $m \times q$ matrix \mathbf{B} and a matrix $\mathbf{C} = \mathbf{AB}$. The dimensions of \mathbf{C} are $n \times q$. Let's define the

entries of these matrices as follows:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \cdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1q} \\ b_{21} & b_{22} & \cdots & b_{2q} \\ \vdots & \cdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mq} \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1q} \\ c_{21} & c_{22} & \cdots & c_{2q} \\ \vdots & \cdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nq} \end{bmatrix}$$

The ij -th entry of \mathbf{C} , denoted by c_{ij} is defined as follows

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj} = \sum_{k=1}^m a_{ik}b_{jk}$$

One way of envisioning matrix multiplication is if we consider the columns of input matrix \mathbf{B} as a set of column vectors, we can operate upon these column vectors one at a time with matrix \mathbf{A} and the resulting vectors will be the corresponding columns of the output matrix \mathbf{C} .

Consider the following matrices:

$$\mathbf{A} = \begin{bmatrix} -2 & 4 \\ 0 & 3 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} 5 & -3 \\ -1 & -1 \end{bmatrix}$$

17. Find the matrix product \mathbf{AB} .

18. Find the matrix product \mathbf{BA}

Note that these two products are NOT equal. In general, matrix multiplication, unlike scalar multiplication, is NOT commutative. In other words, in general $\mathbf{AB} \neq \mathbf{BA}$. However, the distributive property IS valid for matrixes: $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$ so long as we keep the order of the multiplication the same $(\mathbf{B} + \mathbf{C})\mathbf{A} = \mathbf{BA} + \mathbf{CA}$. In order to demonstrate this we need to define matrix addition. Matrix addition is the same as array addition: you simply add the corresponding components together: if $\mathbf{A} + \mathbf{B} = \mathbf{C}$ then $a_{ij} + b_{ij} = c_{ij}$.

In addition to matrices **A** and **B** defined above, consider the matrix

$$\mathbf{C} = \begin{bmatrix} -5 & -1 \\ -3 & 2 \end{bmatrix}$$

19. Calculate $\mathbf{A}(\mathbf{B} + \mathbf{C})$.
20. Calculate $\mathbf{AB} + \mathbf{AC}$. Is it equal to your previous answer?

Now let's consider these principles using the transformation matrixes we worked with above.

21. Hold a closed book in front of you, with the top of the book towards the ceiling ($+z = (0, 0, 1)$ direction) and the front of the book pointed towards you ($+x = (1, 0, 0)$ direction), which leaves the opening side of the book pointing towards your right ($+y = (0, 1, 0)$).

 - (a) Rotate the book by 90 degrees clockwise about the $+x$ axis, then from this position, rotate the book by 90 degrees clockwise about the $+z$ axis. Which direction is the cover of the book facing now?
 - (b) Return to the starting position. Now rotate the book by 90 degrees clockwise about the $+z$ axis, and then from this position, rotate the book by 90 degrees clockwise about the $+x$ axis. Which direction is the cover of the book facing now? Is it the same as in part a?
 - (c) The cover of the book is originally pointed towards $(1, 0, 0)$. Operate on this vector with the appropriate sequence of rotation matrixes from above to reproduce your motions from part a. Do you end up with the correct final cover direction?
 - (d) Operate on the $(1, 0, 0)$ vector with the appropriate sequence of rotation matrixes to reproduce the motions from part b. Do you end up with the correct final cover direction?
 - (e) From either of your answers to part c or part d, try, instead of operating on the $(1, 0, 0)$ vector sequentially with one rotation matrix and then the other, take the product of the two rotation matrixes first, and then operate on the $(1, 0, 0)$ with the resultant matrix. Does this reproduce your answer? What if you take the product the other way?

22. If we want to rotate a vector \mathbf{v} by an angle θ about the z axis to get a vector \mathbf{w} , which matrix \mathbf{R} do we use? Oops, now we changed our mind and decided we didn't want to do that after all. What matrix \mathbf{S} can we multiply \mathbf{w} by to get back to \mathbf{v} ? \mathbf{R} and \mathbf{S} are known as inverses of each other. What happens if we multiple these two matrices together?