# October 18th

We won't get too deep into the theory...

**Definition 1.** *A* ***matrix*** *is a rectangular grid of numbers,* $A = (a_{i,j})_{1\le i\le m, 1\le j\le n}$*, typically written as*

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & \dots & \dots & a_{m,n} \end{pmatrix}$$

*The size of this matrix is* $m \times n$ *(read 'm-by-n', m rows and n columns)*

One can add matrices of the same size: the $(i, j)$-entry of $A+B$ is $a_{i,j}+b_{i,j}$

If $A$ is an $m \times n$ matrix and $B$ is an $n \times p$ matrix (so, the number of columns of $A$ is the number of rows of $B$), then their product $C = AB$ is an $m \times p$ matrix with entries

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$$

In other words, the $(i, j)$-entry is given by the 'dot product' of the $i$-th row of $A$ with the $j$-th column of $B$.

If a matrix is $m \times 1$, we call it a **column vector** (while if it is $1 \times n$, we call it a row vector). Typically we write matrices with capital letters $A, B, C, ...$ and vectors with lower-case letters $v, w, u, ....$

We are concerned only with square matrices, i.e. $n \times n$, and column vectors, $n \times 1$. Notice in this case, the product of a square matrix with a square matrix is another square matrix (all with the same size), and the product of a square matrix with a column vector is another column vector.

I prefer to think of the product of a matrix and a column vector like so: an $n\times n$ matrix is the same as $n$ columns (read left to right), and multiplying this by a column vector $v$ means multiplying the $k$-th column by the $k$-th entry in $v$ to get $n$ column vectors, and adding them all up.

One motivation for introducing matrices is as a notation for systems of equations: the system

$$a_{1,1}x_1 + a_{1,2}x_2 = b_1$$
$$a_{2,1}x_1 + a_{2,2}x_2 = b_2$$

can be written as the 'vector equation'

$$A\vec{x} = \vec{b}$$

For concreteness, the example

$$\begin{aligned} 2x_1 + 1x_2 &= 3 \\ 3x_1 + 5x_2 &= 1 \end{aligned}$$

can be written

$$\begin{pmatrix} 2 & 1 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

Check that $x_1 = 2$ and $x_2 = -1$ solves this system.
Examples: Using the definition of matrix multiplication:

$$\begin{aligned} \begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 3 & -1 \end{pmatrix} &= \begin{pmatrix} 3 \cdot 1 + 2 \cdot 3 & 3 \cdot 4 + 2 \cdot (-1) \\ 0 \cdot 1 + 1 \cdot 3 & 0 \cdot 4 + 1 \cdot (-1) \end{pmatrix} \\ &= \begin{pmatrix} 9 & 10 \\ 3 & -1 \end{pmatrix} \end{aligned}$$

Taking our 'column-centric' perspective, multiplication with a vector:

$$\begin{aligned} \begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 5 \end{pmatrix} &= 2 \cdot \begin{pmatrix} 3 \\ 0 \end{pmatrix} + 5 \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 6 \\ 0 \end{pmatrix} + \begin{pmatrix} 10 \\ 5 \end{pmatrix} \\ &= \begin{pmatrix} 16 \\ 5 \end{pmatrix} \end{aligned}$$

Another example: for some angle $\theta$, consider the matrix

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

If we multiply this with some vectors:

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

and

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

We can draw these, and notice the effect this matrix has is to rotate everything (counter-clockwise) by an angle $\theta$.

We should point out explicitly that

**Proposition 1.** *For any matrix $A$, any vectors $\vec{v}, \vec{w}$, and any real number $k$, we have*

$$A(\vec{v} + \vec{w}) = A\vec{v} + A\vec{w}$$

*and*

$$A(k \cdot \vec{v}) = k \cdot A\vec{v}$$

Going back to our previous example, notice the multiplication

$$\begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$$

so geometrically this vector is stretched but keeps its direction.

Similarly,

$$\begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

so this vector actually stays the same when we apply this matrix multiplication.

On the other hand, the matrix $\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ will not keep any vector's direction the same, since it gives a rotation.

How can we tell whether a matrix $A$ has a vector $v$ which keeps its direction? This means that the vector $v$ just gets stretched, say by a factor $\lambda$. Then

$$A\vec{v} = \lambda\vec{v}$$

or

$$A\vec{v} - \lambda\vec{v} = \vec{0}$$

or

$$(A - \lambda Id)\vec{v} = \vec{0}$$

Obviously $\vec{v} = \vec{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ will satisfy this equation (this is called the 'trivial solution').

It is a fact from linear algebra that

**Theorem 1.** *If $M$ is a matrix, the equation*

$$M\vec{v} = \vec{0}$$

*has a non-trivial solution if and only if* $\det(M) = 0$.

This 'det' thing is the **determinant** of the matrix. For a $2 \times 2$ matrix this is given by the formula

$$\det(\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

(the product of the main diagonal minus the product of the off-diagonal)

In the $3 \times 3$ case, there is again some complicated formula:

$$\det(\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = aei - afh + dhc - dbi + gbf - gec$$

Of course, instead of memorizing a formula for each $n \times n$ case, we should give some conceptual description of what the determinant is.

Geometrically: the determinant of a matrix is the ratio by which an $n$-dimensional box's volume changes after applying the matrix.

Algebraically: the determinant of a matrix is the unique function with input an ordered list of $n$ column vectors which is multilinear (if you write one of the columns as the sum of two vectors, the determinant is the sum of the determinants you get for each vector separately), skew-symmetric (if you switch the order of two columns, you multiply the determinant by $-1$), and evaluates to 1 on the identity matrix.

Let's go back to our question: when does a matrix $A$ leave the direction of some vector $v$ the same? We saw that we need

$$A\vec{v} = \lambda\vec{v}$$

or

$$(A - \lambda Id)\vec{v} = 0$$

So, we must have

$$\det(A - \lambda Id) = 0$$

This is called the **characteristic polynomial** of the matrix $A$: in the $2 \times 2$ case, our formula above means that this is a quadratic polynomial in $\lambda$ (which we can solve). Similarly in the $n \times n$ case we get a degree $n$ polynomial in the variable $\lambda$.

The terminology is

**Definition 2.** *For a matrix $A$, a vector $\vec{v}$ is called an **eigenvector** if*

$$A\vec{v} = \lambda\vec{v}$$

*for some number $\lambda$. A number $\lambda$ such that there is such a vector is called an **eigenvalue** of $A$.*

Example: Consider the matrix $A = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}$. The characteristic polynomial is

$$\begin{aligned}
\det(A - \lambda I) &= \det(\begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}) = 0 \\
&= \det(\begin{pmatrix} 2-\lambda & 3 \\ 1 & 2-\lambda \end{pmatrix}) \\
&= (2-\lambda)(2-\lambda) - 3 \cdot 1 \\
&= 4 - 4\lambda + \lambda^2 - 3 \\
0 &= \lambda^2 - 4\lambda + 1
\end{aligned}$$

so that $\lambda = 2 \pm \sqrt{3}$

If we consider $\lambda = 2 + \sqrt{3}$, then

$$A - \lambda I = \begin{pmatrix} -\sqrt{3} & 3 \\ 1 & -\sqrt{3} \end{pmatrix}$$

and we want a solution to $(A - \lambda I)\vec{v} = \vec{0}$. Notice that the second column is $\sqrt{3}$ times the first, so that

$$\begin{aligned}\begin{pmatrix}-\sqrt{3} & 3\\ 1 & -\sqrt{3}\end{pmatrix}\begin{pmatrix}\sqrt{3}\\ 1\end{pmatrix} &= \sqrt{3}\begin{pmatrix}-\sqrt{3}\\ 1\end{pmatrix} + 1\cdot\begin{pmatrix}3\\ -\sqrt{3}\end{pmatrix}\\ &= \begin{pmatrix}-3\\ \sqrt{3}\end{pmatrix} + \begin{pmatrix}3\\ -\sqrt{3}\end{pmatrix}\\ &= \begin{pmatrix}0\\ 0\end{pmatrix}\end{aligned}$$

and $\vec{v} = \begin{pmatrix}\sqrt{3}\\ 1\end{pmatrix}$ is a solution. Recalling the point of the discussion, this means

$$\begin{pmatrix}2 & 3\\ 1 & 2\end{pmatrix}\begin{pmatrix}\sqrt{3}\\ 1\end{pmatrix} = (2+\sqrt{3})\begin{pmatrix}\sqrt{3}\\ 1\end{pmatrix}$$

so this vector $\begin{pmatrix}\sqrt{3}\\ 1\end{pmatrix}$ points in the same direction after this matrix multiplication.

What about our rotation matrix example? The characteristic polynomial is

$$\begin{aligned}\det\left(\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix} - \lambda\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\right) &= \det\left(\begin{pmatrix}\cos(\theta)-\lambda & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)-\lambda\end{pmatrix}\right)\\ &= (\cos(\theta)-\lambda)(\cos(\theta)-\lambda) + \sin^2(\theta)\\ &= \cos^2(\theta) - 2\cos(\theta)\lambda + \lambda^2 + \sin^2(\theta)\\ &= 1 - 2\cos(\theta)\lambda + \lambda^2\end{aligned}$$

This has solutions

$$\lambda_{\pm} = \frac{2\cos(\theta) \pm \sqrt{4\cos^2(\theta) - 4}}{2}$$

Notice that $\cos^2(\theta) \leq 1$, so this quadratic only has imaginary solutions (except for what case?), which makes sense since we said before that no direction is left unchanged by the rotation.

To be explicit, the method for finding eigenvectors of a matrix $A$ is:

1. Form the characteristic polynomial $\det(A - \lambda Id) = 0$, and solve for $\lambda$. These are the possible eigenvalues of $A$.

2. For each $\lambda$ found in step 1, look at $A - \lambda Id$. Hopefully, the columns are related to each other (in the $2 \times 2$ case, one of the columns should be a number times the other column). Then you can add some combination of the columns to get 0.

3. Thinking of a linear combination of columns as given by multiplication of a matrix with a vector gives a vector $\vec{v}$ so that $(A - \lambda Id)\vec{v} = \vec{0}$, or

$$A\vec{v} = \lambda Id\vec{v} = \lambda\vec{v}$$

   as desired.