18. Powers of matrices

Let's start with three related problems.

Consider the sequence

$$f_0 = 0, 1, 1, 2, 3, 5, 8, 13, \cdots, f_n, \cdots.$$

This sequence satisfies the recurrence

$$f_n = f_{n-2} + f_{n-1}.$$

It is called the Fibonacci sequence. As a motivating question, what is the nth term? That is, can we find a closed form expression for f_n?

For the second problem, suppose we have a physical system, where we keep iterating some action. What is the long term qualitative behaviour of the system?

Here is a seemingly unrelated third problem. Consider the matrix

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

What is A^{100}? Even computing small powers of A looks like a pain.

A much easier problem is to compute powers of a diagonal matrix

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

Then

$$D^2 = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} \qquad \text{and} \qquad D^3 = \begin{pmatrix} 1 & 0 \\ 0 & 8 \end{pmatrix}.$$

In fact it is not hard to see that

$$D^n = \begin{pmatrix} 1 & 0 \\ 0 & 2^n \end{pmatrix}.$$

The idea is to reduce computing powers of A to powers of a diagonal matrix, which is easy.

To see how to do this, let us go back to the problem of computing the nth term f_n of the Fibonnaci sequence. To compute the nth term, we need the previous two terms. This suggests we should create a vector

$$\vec{v}_n = \begin{pmatrix} f_{n-1} \\ f_n \end{pmatrix}.$$

We then have

$$\vec{v}_{n+1} = \begin{pmatrix} f_n \\ f_{n+1} \end{pmatrix} = \begin{pmatrix} f_n \\ f_{n-1} + f_n \end{pmatrix}.$$

The key point is that the last vector is just $A\vec{v}_n$, where

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

In other words $\vec{v}_n = A^{n-1}\vec{v}_1$, where

$$\vec{v}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Now if we have a diagonal matrix and we apply it to a vector, what happens? If we apply the diagonal matrix

$$D = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}.$$

to $\vec{v} = (1, 1)$, we get

$$\begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix}.$$

In general we have

$$D^n\vec{v}_1 = \begin{pmatrix} 1 \\ \frac{1}{2^n} \end{pmatrix}.$$

The key point is that if n is large, then $1/2^n$ is negligible in comparison with 1, so that $D^n\vec{v}_1$ is very close to

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Note that $D\vec{e}_1 = \vec{e}_1$. On the other hand

$$D\vec{e}_2 = \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix} = \frac{1}{2}\vec{e}_2.$$

In fact if D is a diagonal matrix, with entries $\lambda_1, \lambda_2, \dots, \lambda_n$ on the main diagonal, then we have $D\vec{e}_i = \lambda_i\vec{e}_i$. This motivates:

Definition 18.1. *Let A be an $n \times n$ matrix. We say that $\vec{v} \neq 0$ is an eigenvector with eigenvalue λ if $A\vec{v} = \lambda\vec{v}$.*

So, a diagonal matrix D, with diagonal entries $\lambda_1, \lambda_2, \dots, \lambda_n$, has eigenvectors $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$, with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$. Note that the eigenvectors are a basis for $\mathbb{R}^n$.

One can push this a little bit further. We say that a square matrix is upper triangular if every entry below the main diagonal is zero. If A is upper triangular then the eigenvalues of A are the entries on the main diagonal.

Example 18.2. *The eigenvalues of*

$$\begin{pmatrix} -2 & 10 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & 7 \end{pmatrix}$$

are -2, 3 and 7.

There are couple of other cases when one can figure out directly the eigenvalues and eigenvectors. Consider the linear function

$$f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \qquad \text{given by} \qquad (x, y) \longrightarrow (x, -y).$$

We are looking for invariant lines. We have already seen this function. It represents reflection in the x-axis. So the whole x-axis is fixed. Thus $\vec{e}_1 = (1, 0)$ is an eigenvector with eigenvalue 1. The y-axis is fixed but it is also flipped. $\vec{e}_2 = (1, 0)$ is an eigenvector with eigenvalue -1.

What about reflection in another line? If the line is spanned by the vector $\vec{v}$ then $\vec{v}$ is an eigenvector with eigenvalue 1. If $\vec{w}$ is a vector which makes a right angle to $\vec{v}$ then $\vec{w}$ is flipped. It is an eigenvector with eigenvalue -1.

For example, consider reflection in the line $y = x$.

$$f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \qquad \text{given by} \qquad (x, y) \longrightarrow (y, x).$$

Then $(1, 1)$ is an eigenvector with eigenvalue 1 and $(1, -1)$ is an eigenvector with eigenvalue -1.

What about rotation? If we rotate around the origin in $\mathbb{R}^2$ then most of the time there are no eigenvectors, unless we rotate through exactly π. In this case any vector is an eigenvector with eigenvalue -1.

If we perform a rotation in space then there is always an axis of rotation. Any vector which spans the axis is an eigenvector with eigenvalue 1.

In all these cases eigenvectors with different eigenvalues are independent and eigenvectors with the same eigenvalue are a linear subspace.

Theorem 18.3. *Let A be an $n \times n$ matrix and let $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ be eigenvectors of A with distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_k$.*

Then $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ are independent. In particular if $k = n$ then $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ are a basis of eigenvectors for $\mathbb{R}^n$.

Proof. Suppose not. Suppose that $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ are dependent. We will derive a contradiction. By assumption there are scalars $r_1, r_2, \dots, r_k$, not all zero, such that

$$\vec{0} = r_1\vec{v}_1 + r_2\vec{v}_2 + \cdots + r_k\vec{v}_k.$$

We suppose that k is minimal with this property. In particular we may assume that $r_i \neq 0$ for all i. Clearly $k > 1$. We apply A to both sides of the equation above. We get

$$\begin{aligned}\vec{0} &= A \cdot \vec{0} \\ &= A(r_1\vec{v}_1 + r_2\vec{v}_2 + \cdots + r_k\vec{v}_k) \\ &= r_1 A\vec{v}_1 + r_2 A\vec{v}_2 + \cdots + r_k A\vec{v}_k \\ &= r_1\lambda_1\vec{v}_1 + r_2\lambda_2\vec{v}_2 + \cdots + r_k\lambda_k\vec{v}_k.\end{aligned}$$

Take the first equation and multiply by λ_k. We get

$$\vec{0} = r_1\lambda_k\vec{v}_1 + r_2\lambda_k\vec{v}_2 + \cdots + r_k\lambda_k\vec{v}_k$$

$$\vec{0} = r_1\lambda_1\vec{v}_1 + r_2\lambda_2\vec{v}_2 + \cdots + r_k\lambda_k\vec{v}_k.$$

We subtract the second equation from the first equation:

$$\vec{0} = r_1(\lambda_k - \lambda_1)\vec{v}_1 + r_2(\lambda_k - \lambda_2)\vec{v}_2 + \cdots + r_{k-1}(\lambda_k - \lambda_{k-1})\vec{v}_{k-1}.$$

Now $s_i = r_i(\lambda_k - \lambda_i) \neq 0$, since the eigenvalues are distinct. But then we found a linear dependence involving fewer eigenvectors. This contradicts our choice of k. The only possibility is that the eigenvectors are independent to start with. □

How does one compute the eigenvalues and eigenvectors? Well if λ is an eigenvalue then

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

Note that

$$\lambda\vec{v} = \lambda I_n\vec{v}.$$

Rearranging we get

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

So an eigenvector $\vec{v}$ of A with eigenvalue λ is the same as an element of the nullspace of $B = A - \lambda I_n$. The set of all eigenvectors with eigenvalue λ,

$$\{\, \vec{v} \in \mathbb{R}^n \mid A\vec{v} = \lambda\vec{v} \,\}$$

is a linear subspace of $\mathbb{R}^n$, called an eigenspace.