Exponential flows

The solutions of ordinary differential equations can be seen as “flows” or 1-parameter groups. We now study the simplest such groups and the homogeneous linear ordinary differential equations with constant coefficients that these solve.

The simplest 1-parameter groups

Scaling and shrinking

A one-parameter group is a map $f : \mathbb{R} \to G$, where G is a group which we write multiplicatively. In other words, it “turns” addition into multiplication. The simplest such map that we know is “taking the power of”. This is more mathematically defined using the exponential function:

$$\exp(t) = 1 + t + \frac{t^2}{2!} + \cdots = \sum_{k=0}^{\infty} \frac{t^k}{k!}$$

The fundamental identity $\exp(a+b) = \exp(a)\cdot\exp(b)$ can be obtained by checking it “term by term” where it becomes,

$$\frac{(a+b)^n}{k!} = \sum_{r=0}^{k} \frac{a^r}{r!} \cdot \frac{b^{k-r}}{(k-r)!}$$

which is just another way to write the Binomial theorem! Note that this checking term-by-term is valid since the above series is *abosolutely convergent*. Thus we see that exp converts addition into multiplication. More generally, for any constant c we have $\exp((t+s)c) = \exp(tc) \cdot \exp(sc)$. In particular, we can take $c = \log(2)$ in which case $\exp(c) = 2$ so that we can think of this as $2^{t+s} = 2^t \cdot 2^s$.

Rotation

Another “standard” 1-parameter group is given by rotations in the plane. Clearly rotation by an angle t followed by a rotation by a rotation by the angle s is the same as rotation by the angle $t + s$. Rotation by the angle t is the linear transformation associated the matrix:

$$\begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

Thus, the required identity is:

$$\begin{pmatrix} \cos(t+s) & -\sin(t+s) \\ \sin(t+s) & \cos(t+s) \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \cdot \begin{pmatrix} \cos s & -\sin s \\ \sin s & \cos s \end{pmatrix}$$

This identity is just a re-statement of the “law of addition for sine and cosine”.

Shearing

Yet another simple matrix identity is

$$\begin{pmatrix} 1 & (t+s) \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}$$

This corresponds to the flow in the plane given by $(x, y) \mapsto (x + yt, y)$ which is called a "shearing"; each point moves to the right an amount proportional to its height above the x-axis.

Exponential

As seen above, the exponential function $\exp : \mathbb{R} \to \mathbb{R}^{>0}$ is given in terms of the power series

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

What this means is that if we define the *polynomial* functions

$$e_N(x) = \sum_{k=0}^{N} \frac{x^k}{k!}$$

Then, for every x, the sequence of real numbers $e_N(x)$ converges to a real number $\exp(x)$. In fact, for each *positive* real number x, the sequence of numbers $e_N(x)$ is *increasing* and this is a bounded sequence. Hence, by the least upper bound principle it converges to a real number and we define that number as $\exp(x)$. Moreover, if we fix any positive real number B, then given any approximation error $\epsilon > 0$, we can find a fixed N so that $e_m(x)$ is within ϵ of the actual value $\exp(x)$ for *all* x with $|x| < B$ and for *all* $m \geq N$. This is the best possible kind of convergence since, if we know how large the x's are which we are working with, we can calculate $\exp(x)$ to any chosen level of accuracy with a fixed formula given by $e_N(x)$.

In a number of situations, matrices can be seen as a generalisation of the concept of number. Thus, for an $n \times n$ matrix A, we can try to make sense of the "infinite sum":

$$\exp(A) = \mathbf{1}_n + A + (1/2)A^2 + (1/6)A^3 + \cdots = \sum_{k=0}^{\infty} (1/k!)A^k$$

where $\mathbf{1}$ denotes the $n \times n$ identity matrix. To understand this, we need to understand the notion of convergence of matrices. The *finite* sums

$$e_N(A) = \mathbf{1}_n + A + (1/2)A^2 + (1/6)A^3 + \ldots (1/N!)A^N$$

certainly make sense and give us an $n \times n$ matrix $e_N(A)$. To take the limit of these as N goes to infinity, the natural sense is that *each entry* of these matrices converges as N goes to infinity. Note that the (i, j)-th entry of $e_N(A)$ depends on various entries of A not *just* the (i, j)-th one! So if we try to write a direct formula for this entry it is likely to be very complicated!

Since the matrix A has n^2 entries we can find a common positive constant L so that, for all (i, j), the (i, j)-th entry $A_{i,j}$ of A is bounded by L; in other words, $|A_{i,j}| \leq L$. Recall the matrix multiplication formula:

$$(B \cdot C)_{i,j} = B_{i,1}C_{1,j} + B_{i,2}C_{2,j} + \cdots + B_{i,n}C_{n,j}$$

Exercise: Using the above formula, prove by induction that

$$|(A^k)_{i,j}| \leq n^{k-1}L^k \leq n^k L^k$$

(since $n \geq 1$).

It follows easily that

$$|e_N(A)_{i,j}| \leq e_N(nL)$$

By the standard theory of convergent sequences one can now conclude that $e_N(A)_{i,j}$ converges for each (i, j). In conclusion, we see that $\exp(A)$ makes sense for any matrix A.

Given a pair of matrices A and B, we would like to understand how $\exp(A + B)$ can be expressed in terms of $\exp(A)$ and $\exp(B)$. Here, the analogy with numbers breaks down! To apply the Binomial theorem to $(A + B)^k$ we need $A \cdot B = B \cdot A$ which does not work with matrices in general. However, *if* this commutativity condition holds then, as above we can check (using absolute convergence again) that

$$\exp(A + B) = \exp(A) \cdot \exp(B) \text{ if } A \cdot B = B \cdot A$$

In particular, we note that $(tA) \cdot (sA) = stA^2$ for s and t scalars, so that

$$\exp((t + s)A) = \exp(tA + sA) = \exp(tA) \cdot \exp(sA)$$

Thus, every matrix A gives rise to a 1-parameter flow! Let us look at some simple examples.

Some examples

Taking A to be $\mathbf{1}_n$, the identity matrix, we see that the diagonal terms of $\exp(t\mathbf{1}_n)$ are given by $\exp(t)$ and the off diagonal terms are all 0. So $\exp(t\mathbf{1}_n) = \exp(t)\mathbf{1}_n$[1]. More generally, if we take A to be a diagonal matrix D with the numbers $c_1, \ldots, c_n$ on the diagonal, we see easily that $\exp(tD)$ is the diagonal matrix with the

[1] This should *not* give the impression that $\exp(tA) = \exp(t)\exp(A)$ for a general matrix! It is easy to write examples of A where this is false!

functions $\exp(tc_1), \ldots, \exp(tc_n)$ on the diagonal. In other words, we have the linear operation that scales (or shrinks!) at different rates along different axes.

Next, let us take the case where A is the 2×2 matrix $I = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. We note that $I^2 = -\mathbf{1}_2$. Hence, it follows that

$$I^k = \begin{cases} (-1)^r \mathbf{1}_2 & ; k = 2r \\ (-1)^r I & ; k = 2r+1 \end{cases}$$

It then follows that

$$\exp(tI) = \begin{pmatrix} f(t) & -g(t) \\ g(t) & f(t) \end{pmatrix}$$

where $f(t)$ and $g(t)$ are given by the power series as below

$$f(t) = \sum_{r=0}^{\infty} (-1)^r \frac{t^{2r}}{(2r)!}$$

$$g(t) = \sum_{r=0}^{\infty} (-1)^r \frac{t^{2r+1}}{(2r+1)!}$$

We recognise that $f(t)$ is the power series for $\cos t$ and $g(t)$ is the power series for $\sin t$. In other words, $\exp(tI)$ is the familiar 1-parameter group of rotations. We further note that

$$\exp(tI) = \cos t \mathbf{1}_2 + \sin t I$$

which is the matrix version of de Moivre's identity!

Finally, let us consider the matrix $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. In this case $N^2 = 0$, so $\exp(tN) = \mathbf{1}_2 + tN$. Writing this out in full we have

$$\exp(tN) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$

This is the 1-parameter group of shearing as seen earlier.

It is not a coincidence that the three examples we looked at are the scaling, rotation and shear. As we shall see below, *all* flows of the type $\exp(tA)$ are made up of such flows.

Homogeneous Linear Ordinary Differential Equations with Constant Coefficients

The matrix entries of $\exp(tA)$ are differentiable functions of t. In fact, we check

$$\begin{aligned}\frac{d}{dt}\exp(tA) &= \lim_{h\to 0}\frac{\exp((t+h)A) - \exp(tA)}{h}\\ &= \lim_{h\to 0}\frac{\exp(hA)\cdot\exp(tA) - \exp(tA)}{h}\\ &= \lim_{h\to 0}\frac{\exp(hA) - \mathbf{1}_n}{h}\cdot\exp(tA)\\ &= A\cdot\exp(tA)\end{aligned}$$

where we have used the power series for $\exp(hA)$ to check that $\frac{\exp(hA)-\mathbf{1}_n}{h}$ goes to A as h goes to 0.

In particular, if $\vec{v_0}$ is *any* $n\times 1$ column vector and we define $\vec{v}(t) = \exp(tA)\cdot\vec{v}_0$, then:

$$\frac{d\vec{v}}{dt} = A\cdot\vec{v}$$

Turning this around, given *any* equation of the above form, we have written down the associated flow and the solution given a starting point $\vec{v}_0$. Since the right-hand side of this equation depends *linearly* on the entries of $\vec{v}$ and the coefficients (which are the entries $A_{i,j}$) are given constants, we say that this is a linear differential equation with constant coefficients. Since the multiple of a solution is also a solution, we say that this is a *homogeneous* equation. While we have actually determined the complete flow, it is worth noting that given the "initial condition" $\vec{v}(0) = \vec{v}_0$, we have found the solution which satisfies this condition. Thus we have solved an *initial value problem*.

We will see a general result that will imply that the solution to this problem (and more general ones) is *unique*. Hence, we have indeed found *the* solution to the initial value problem (IVP) for a homogeneous linear Ordinary Differential Equation (ODE). This solution plays an important role in understanding more general equations, so we will spend a little more time with it!