General linear ordinary differential equation

A linear ordinary differential equation has the form $(d\vec{v}/dt) = A \cdot \vec{v}$ where A is an $n \times n$ matrix and $\vec{v}$ is an column n-vector, where A is *independent* of $\vec{v}$. So far we have studied the case where A does not vary with t. We now want to study the more general case where $A = A(t)$ may vary with t.

Before we begin, here is a clarificatory remark. In the introduction to this course, we had pointed out that a "time-dependent" equation $(d\vec{v}/dt) = \vec{f}(t, \vec{v})$ can be converted into a time-independent one by introducing $\vec{w}$ as a column vector with $(n+1)$ entries; the first n entries are the same as $\vec{v}$ and the last entry is t. On the right hand side we can then substitute w_{n+1} in place of t to write $(d\vec{w}/dt) = \vec{g}(\vec{w})$ where the first n entries of $\vec{g}$ are $f(w_{n+1}, \vec{w})$ and the last entry is 1 $(= (dw_{n+1}/dt))$. If we attempt to use this in the above case, we see that the new matrix A becomes dependent on w_{n+1} and hence the equation is no longer linear.

To summarise, it is *important* that A remains *independent* of $\vec{v}$ for us to call it a *linear* equation. However, we *can* allow A to depend on t as we will now do.

Linearity of the solution

As a consequence of the general statement of existence and uniqueness of solutions of ordinary differential equation, we will see that under some *very mild* hypothesis on A (such as continuity), there is a unique function $\vec{v}(t)$ which satisfies

$$\frac{d\vec{v}}{dt} = A(t) \cdot \vec{v} \text{ and } \vec{v}(0) = \vec{v_0}$$

for any given *initial* vector $\vec{v_0}$. This solution exists for small enough interval around 0 (for t) which is *independent* of the chosen $\vec{v_0}$.

In fact, if we start at a different initial time t_0 and put the condition $\vec{v}(t_0) = \vec{v_0}$ instead of the above initial condition, then there is a unique solution $\vec{v}(t)$ for a small enough interval around t_0.

Let $\vec{u_0} = a\vec{v_0} + \vec{w_0}$ for some *constant* a and some chosen *fixed* vectors $\vec{v_0}$ and $\vec{w_0}$. We use the notation $\vec{u}$ to denote the solution of

$$\frac{d\vec{u}}{dt} = A(t) \cdot \vec{u} \text{ and } \vec{u}(0) = \vec{u_0}$$

and we use the notation $\vec{w}$ to denote the solution of

$$\frac{d\vec{w}}{dt} = A(t) \cdot \vec{w} \text{ and } \vec{w}(0) = \vec{w_0}$$

Consider the vector valued function $\vec{u'} = a\vec{v} + \vec{w}$. We note that

$$\vec{u'}(0) = a\vec{v}(0) + \vec{w}(0) = a\vec{v_0} + \vec{w_0} = \vec{u_0}$$

Moreover, since a is a constant

$$\frac{d\vec{u'}}{dt} = a\frac{d\vec{v}}{dt} + \frac{d\vec{w}}{dt} = aA(t)\cdot\vec{v} + A(t)\cdot\vec{w}$$

Now, matrix multiplication *commutes* with scalar multiplication (by a) and *preserves addition*. So

$$aA(t)\cdot\vec{v} + A(t)\cdot\vec{w} = A(t)\cdot a\vec{v} + A(t)\cdot\vec{w} = A(t)\cdot(a\vec{v} + \cdot\vec{w}) = A\vec{u'}$$

In other words, we see that $\vec{u'}$ *also* satisfies

$$\frac{d\vec{u'}}{dt} = A(t)\cdot\vec{u'} \text{ and } \vec{u'}(0) = \vec{u_0}$$

By the *uniqueness* of the solution, we see that $\vec{u'} = \vec{u}$.

We have shown the following. If $\vec{v}$ is the solution of the Linear ODE with initial value $\vec{v_0}$ and $\vec{w}$ is the solution of the same Linear ODE with (the other) initial value $\vec{w_0}$, then $a\vec{v}+\vec{w}$ is the solution of the same Linear ODE with the combined initial value $a\vec{v_0} + \vec{w_0}$.

Before proceeding further, let us see how linearity is important. Consider a general ODE $(d\vec{v}/dt) = \vec{f}(t,\vec{v})$ and let $\vec{v}$ be a solution of this with initial value $\vec{v_0}$ and $\vec{w}$ be a solution of this with initial value $\vec{w_0}$. We can form $\vec{u'} = a\vec{v} + \vec{w}$ as we did above. What does it satisfy? It is clear that, as above, we have

$$\vec{u'}(0) = a\vec{v}(0) + \vec{w}(0) = a\vec{v_0} + \vec{w_0}$$

However, when we examine the derivative we see

$$\frac{d\vec{u'}}{dt} = a\frac{d\vec{v}}{dt} + \frac{d\vec{w}}{dt} = a\vec{f}(t,\cdot\vec{v}) + \vec{f}(t,\cdot\vec{w})$$

The expression on the right-hand side *does not*, in general, simplify to $f(t, a\vec{v}+\vec{w})$ since we do not know that $\vec{f}(t,\vec{v})$ is linear with respect to the second (vectorial) argument $\vec{v}$. Thus, we appear to have a solution of an entirely different ordinary differential equation (if at all!).

Linear independence

Given that $\vec{v_i}(t)$, for $i = 1,\dots,k$ are solutions of the linear ordinary differential equation $(d\vec{v}/dt) = A\cdot\vec{v}$. Further, let us assume that $\vec{v_1}(0), \vec{v_2}(0), \dots, \vec{v_k}(0)$ are *linearly independent*. We now claim that, for every t, the vectors $\vec{v_1}(t), \vec{v_2}(t), \dots, \vec{v_k}(t)$ are linearly independent. Suppose that $a_1, a_2, \dots, a_k$ are constants so that

$$a_1\vec{v_1}(t_0) + a_2\vec{v_2}(t_0) + \cdots + a_k\vec{v_k}(t_0) = 0 \text{ for some } t_0$$

By uniqueness of the solution with initial time t_0, it follows that

$$a_1\vec{v_1}(t) + a_2\vec{v_2}(t) + \cdots + a_k\vec{v_k}(t) = 0 \text{ for all } t$$

In particular, it follows that

$$a_1\vec{v_1}(0) + a_2\vec{v_2}(0) + \cdots + a_k\vec{v_k}(0) = 0$$

However, these vectors are given to be linearly independent and so we conclude that $a_i = 0$ for $i = 1, \ldots, k$ as required.

This has a practical consequence. For each i from 1 to n, we let $\vec{v_i}$ denote the solution of the Linear ODE as above with the initial condition $\vec{v_i}(0) = \vec{e_i}$ where

$$\vec{e_i} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

is the standard basis vector of $\mathbb{R}^n$ with 1 in the i-th row and 0 everywhere else. We now define the matrix $G = [v_1, \ldots, v_n]$ whose columns are $\vec{v_i}$; this is a matrix valued function of t. Then, the solution of the initial value problem with initial condition $\vec{v}(0) = \vec{v_0}$ is given by $\vec{v} = G \cdot \vec{v_0}$.

In other words, G plays the *same role* as $\exp(tA)$ did in the case when A has constant coefficients!

Another way to look at this is to say that we only need to find n solutions to the Linear ODE, corresponding to the initial positions $\vec{e_i}$. All other solutions are obtained as linear combinations of these solutions.

Determinant

Given an $n \times n$ matrix G, recall that its determinant is defined inductively as

$$\det(G) = G_{1,1} \det H(1,1) - G_{1,2} \det H(1,2) + \cdots + (-1)^n G_{1,n} \det H(1,n)$$

where $G_{i,j}$ is the (i,j)-th entry of G and $H(i,j)$ is the $(n-1) \times (n-1)$ matrix obtained from G by deleting the i-th row and j-th column. By an application of Leibniz rule and induction we can show that

$$\frac{d \det(G)}{dt} = \det G^{(1)} + \det G^{(2)} + \cdots + \det G^{(n)}$$

where $G^{(k)}$ is the matrix obtained from G by differentiating the k-th column of G.

Now, let us assume that G is the matrix solution to a linear ODE as constructed above. Then each column $\vec{v_i}$ of G satisfies $(d\vec{v_i}/dt) = A \cdot \vec{v_i}$. So the right-hand side in the above expression becomes

$$\det[A \cdot \vec{v_1}, \vec{v_2}, \dots, \vec{v_n}] + \det[\vec{v_1}, A \cdot \vec{v_2}, \dots, \vec{v_n}] + \det[\vec{v_1}, \vec{v_2}, \dots, A \cdot \vec{v_n}]$$

An elementary but instructive exercise can be used to show that this expression is $\mathrm{Tr}(A)\det(G)$ where $\mathrm{Tr}(A)$ denotes the trace of the matrix A. In other words, we have obtained the identity

$$\frac{d\det(G)}{dt} = \mathrm{Tr}(A)\det(G)$$

So, $\det(G)$ satisfies *this* ordinary differential equation. As we saw in the introduction, the solution of this differential equation is

$$\det(G) = \exp\left(\int_0^t \mathrm{Tr}(A)(t)dt\right)$$

where we have used $\det(G)(0) = 1$ since we started with the standard basis at $t = 0$.

We note, in particular, that if the trace of A is *identically* 0, then $\det(G)$ is identically 1; in other words, the flow is "volume preserving".

There are other examples of polynomial identities that can be deduced for G based on certain linear properties of A. This is the starting point of the theory of Lie Algebras. We only mention one important property. If A is a skew-symmetric matrix, then G is orthogonal, i.e. $G \cdot G^t = \mathbf{1}_n$. In other words, if A is skew-symmetric then the flow preserves lengths and angles.

Wronskian

A particular case of the above analysis is when we solve a linear equation of order n with one indeterminate:

$$\frac{d^n x}{dt^n} + a_1(t)\frac{d^{n-1}x}{dt^{n-1}} + \cdots + +a_{n-1}(t)\frac{dx}{dt} + a_n x = 0$$

As mentioned in the introduction, we can introduce new variables $x_1 = x$, $x_2 = (dx/dt)$, and so on till $x_n = (d^{n-1}x/dt^{n-1})$. This equation becomes

$$\frac{d}{dt}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_n & -a_{n-1} & -a_{n-3} & \dots & -a_1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix}$$

It is clear that the trace of the matrix is $-a_1$ in this case. Suppose that $x_i(t)$ denotes the solution with initial condition

$$\frac{d^k x_i}{dt^k} |_{t=0} = \begin{cases} 0 & k \neq i \\ 1 & k = i \end{cases}$$

for $k = 0, \ldots, (N-1)$. Then the Wronskian is defined as

$$W(t) = \det \begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ \frac{dx_1}{dt} & \frac{dx_2}{dt} & \ldots & \frac{dx_n}{dt} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{d^{n-1}x_1}{dt^{n-1}} & \frac{d^{n-1}x_2}{dt^{n-1}} & \ldots & \frac{d^{n-1}x_n}{dt^{n-1}} \end{pmatrix}$$

This is non-zero for all t and is given by $W(t) = \exp\left(\int -a_1(t)dt\right)$. In fact, one can show that for *any* collection of functions $x_1, x_1, \ldots, x_n$ which are linearly independent, the Wronskian as defined above is non-zero.