LECTURE 8

Date of Lecture: January 27, 2021

The symbol ☡ is for flagging a cautionary comment or a tricky argument. It occurs in the margins and is Knuth's version of Bourbaki's "dangerous bend symbol".

An n-tuple $(x_1, \ldots, x_n)$ of symbols (x_i not necessarily real or complex numbers) will also be written as a column vector when convenient. Thus

$$(x_1, \ldots, x_n) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

A map $\boldsymbol{f}$ from a set S to a product set $T_1 \times \cdots \times T_n$ will often be written as an n-tuple $\boldsymbol{f} = (f_1, \ldots, f_n)$, with f_i a map from S to T_i, and hence, by the above convention, as a column vector

$$\boldsymbol{f} = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}.$$

(See Remark 2.2.2 of Lecture 5 of ANA2.)

The default norm on Euclidean spaces of the form $\mathbf{R}^n$ is the Euclidean norm $\| \, \|_2$ and we will simply denote it as $\| \, \|$. The space of $\mathbf{R}$-linear transformations from $\mathbf{R}^n$ to $\mathbf{R}^m$ will be denoted $\mathrm{Hom}_{\mathbf{R}}(\mathbf{R}^n, \mathbf{R}^m)$ and will be identified in the standard way with the space of $m \times n$ real matrices $M_{m,n}(\mathbf{R})$ and the operator norm[1] on both spaces will be denoted $\| \, \|_\circ$. If $m = n$, we write $M_n(\mathbf{R})$ for $M_{m,n}(\mathbf{R})$.

☡ Note that $(x_1, \ldots, x_n) \neq [x_1 \ldots x_n]$. Each side is the transpose of the other.

1. Linear Differential Equations

Let $I \subset \mathbf{R}$ be an interval (closed, open, half-open, but with non-empty interior). Recall that a (vector valued) *linear differential equation* is a differential equation of the form

$$\dot{\boldsymbol{x}}(t) = A(t)\boldsymbol{x}(t) + \boldsymbol{g}(t) \tag{1}$$

where $A\colon I \to M_n(\mathbf{R})$ and $\boldsymbol{g}\colon I \to \mathbf{R}^n$ are continuous maps. Note that the extended phase space is $I \times \mathbf{R}^n$ and (1) is of the form $\dot{\boldsymbol{x}} = \boldsymbol{v}(t, \boldsymbol{x})$ where $\boldsymbol{v}\colon I \times \mathbf{R}^n \to \mathbf{R}^n$ is given by the formula $\boldsymbol{v}(t, \boldsymbol{x}) = A(t)\boldsymbol{x}(t) + \boldsymbol{g}(t)$ for $t \in I$ and $\boldsymbol{x} \in \mathbf{R}^n$. It is easy to see, and this was shown in the proof of Theorem 1.1.1 of Lecture 7, that $\boldsymbol{v}$ is locally Lipschitz in the second variable. In fact it was shown there that $\boldsymbol{v}$ is uniformly Lipschitz on any rectangle of the form $J \times \mathbf{R}^n$ where J is closed (i.e. compact) interval in I.

[1]See §§2.1 of Lecture 5 of ANA2.

If in the above, $\boldsymbol{g}(t) = 0$ for every $t \in I$, then (1) is said to be a *homogeneous linear differential equation.* Explicitly, a homogeneous differential equation on I is an equation of the form

$$\dot{\boldsymbol{x}}(t) = A(t)\boldsymbol{x}(t) \tag{2}$$

where $A\colon I \to M_n(\mathbf{R})$ is a continuous function. Here $\boldsymbol{v}(t, \boldsymbol{x}) = A(t)\boldsymbol{x}$ for $t \in I$ and $\boldsymbol{x} \in \mathbf{R}^n$. In Theorem 1.1.1 of Lecture 7 we showed that with A and $\boldsymbol{g}$ as above, and $(\tau, \boldsymbol{a})$ a fixed point in $I \times \mathbf{R}^n$, the IVP $\dot{\boldsymbol{x}}(t) = \boldsymbol{v}(t, \boldsymbol{x})$, $\boldsymbol{x}(\tau) = \boldsymbol{a}$ has unique solution on all of I.

Assumption. For the rest of this lecture we fix I and A as above. In particular, A is continuous.

1.1. For a non-negative integer k, let $\mathscr{C}^k(I)$ be the short form of $\mathscr{C}^k(I, \mathbf{R}^n)$, the $\mathbf{R}$-vector space of $\mathbf{R}^n$-valued $\mathscr{C}^k$ functions on I (with the usual convention of using one-sided derivatives in the event a boundary point of I lies in I). A reminder: $\mathscr{C}^0(I)$ is the space of continuous $\mathbf{R}^n$-valued functions on I. Consider the map

$$\begin{aligned} \mathscr{C}^1(I) &\xrightarrow{T} \mathscr{C}^0(I) \\ \boldsymbol{f} &\longmapsto \dot{\boldsymbol{f}} - A\boldsymbol{f} \end{aligned} \tag{1.1.1}$$

It is clear that T is a linear transformation. Let

$$S := \ker T. \tag{1.1.2}$$

It is clear that S is the set of solutions of the homogeneous differential equation (2). In particular, S is an $\mathbf{R}$-vector space in a natural way. It is a subspace of $\mathscr{C}^1(I)$. As in Problem 5 of HW2 we see that

Theorem 1.1.3. *Let S and T be as above.*

(a) *The map $T\colon \mathscr{C}^1(I) \to \mathscr{C}^0(I)$ is surjective.*
(b) *The vector space S is n-dimensional.*

Proof. Given $\boldsymbol{g} \in \mathscr{C}^0(I)$, we know that the differential equation (1) has solutions on I by Theorem 1.1.1 of Lecture 7. Indeed fix $t_\circ \in I$ and $\boldsymbol{a} \in \mathbf{R}^n$, and we can find a solution $\boldsymbol{\varphi}$ with the requirement that $\boldsymbol{\varphi}(t_\circ) = \boldsymbol{a}$. Now, $\boldsymbol{\varphi} \in \mathscr{C}^{(}I)$ and $T\boldsymbol{\varphi} = \boldsymbol{g}$. Thus T is surjective. This proves (a).

We now prove (b). Fix $t_\circ \in I$ and let

$$E = E_{t_\circ} \colon S \longrightarrow \mathbf{R}^n$$

be the evaluation map given by

$$E\boldsymbol{f} = \boldsymbol{f}(t_\circ).$$

E is clearly a linear transformation. If $\boldsymbol{f} \in S$ is such that $E(\boldsymbol{f}) = \mathbf{0}$, then $\boldsymbol{f}$ is a solution to the IVP

$$\dot{\boldsymbol{x}} = A\boldsymbol{x}, \quad \boldsymbol{x}(t_\circ) = \mathbf{0}.$$

On the other hand, so is the constant function $\mathbf{0}$. By uniqueness of solutions to IVPs, we see that $\boldsymbol{f} = \mathbf{0}$. Thus E is an injective linear transformation.

Next suppose $\boldsymbol{a} \in \mathbf{R}^n$. We claim $\boldsymbol{a} = E\boldsymbol{f}$ for some $\boldsymbol{f} \in S$. Consider the IVP

$$\dot{\boldsymbol{x}} = A\boldsymbol{x}, \quad \boldsymbol{x}(t_\circ) = \boldsymbol{a}.$$

This IVP has a unique solution $\boldsymbol{f}$ on I. Now, $\boldsymbol{f} \in S$ and $\boldsymbol{f}(t_\circ) = \boldsymbol{a}$. In other words $\boldsymbol{f} \in S$ and $E\boldsymbol{f} = \boldsymbol{a}$. Thus E is surjective and hence we have

$$E \colon S \xrightarrow{\sim} \mathbf{R}^n.$$

This proves that S is n-dimensional. □

Since T is surjective, if $\boldsymbol{g} \in \mathscr{C}^0(I)$, then $T^{-1}(\boldsymbol{g}) \neq \emptyset$. This means, since $S = \ker T$, that $T^{-1}(\boldsymbol{g})$ is a coset of S. In fact if $\boldsymbol{\psi} \in T^{-1}(\boldsymbol{g})$, we have

$$T^{-1}(\boldsymbol{g}) = S + \boldsymbol{\psi}. \tag{1.1.4}$$

We emphasise that (1.1.4) remains true whichever element $\boldsymbol{\psi}$ of $T^{-1}(\boldsymbol{g})$ we happen to pick.

1.1.5. General solutions, particular solutions, and complementary solutions. Let $\boldsymbol{\varphi}_1, \dots, \boldsymbol{\varphi}_n$ be linearly independent elements of S. Since $\dim_{\mathbf{R}} S = n$, this is equivalent to saying $\boldsymbol{\varphi}_1, \dots, \boldsymbol{\varphi}_n$ is a basis of S. An arbitrary element of S can then be expressed as

$$\boldsymbol{\varphi} = c_1\boldsymbol{\varphi}_1 + \cdots + c_n\boldsymbol{\varphi}_n \tag{1.1.5.1}$$

where $c_1, \dots, c_n$ are arbitrary constants. The expression in (1.1.5.1) (with arbitrary constants c_i, $i = 1, \dots, n$) is called the *general solution to the homogeneous equation* (2).

Let $\boldsymbol{g} \in \mathscr{C}^0(I)$ and consider the equation (1). Pick an element $\boldsymbol{\psi}_p$ of $T^{-1}(\boldsymbol{g})$. Note that $\boldsymbol{\psi}_p$ is a solution of (1). In view of (1.1.4), we see that by varying $(c_1, \dots, c_n)$ in $\mathbf{R}^n$, the expression

$$\boldsymbol{\psi} = c_1\boldsymbol{\varphi}_1 + \cdots + c_n\boldsymbol{\varphi}_n + \boldsymbol{\psi}_p \tag{1.1.5.2}$$

gives all solutions of (1). For a fixed solution $\boldsymbol{\psi}_p$ of (1) the correspondence between $\boldsymbol{c} = (c_1, \dots, c_n) \in \mathbf{R}^n$ and $\boldsymbol{\psi}$ is bijective. In the expression for $\boldsymbol{\psi}$ in (1.1.5.2), $\boldsymbol{\psi}_p$ is called a *particular solution of* (1). The expression in (1.1.5.2), with $c_1, \dots, c_n$ arbitrary constants and $\boldsymbol{\psi}_p$ a fixed particular solution, is called the *general solution to* (1).

The general solution to (2) (i.e. expression in (1.1.5.1)) is also called *the complementary solution to the (possibly) inhomogeneous equation* (1). It is not a solution of (1) unless $\boldsymbol{g} = \boldsymbol{0}$, so the terminology can be confusing.

2. Variation of Parameters

In this section, in addition fixing the continuous function $A \colon I \to M_n(\mathbf{R})$, we also fix a continuous function $\boldsymbol{g} \colon I \to \mathbf{R}^n$. Equation (1), when invoked, will be with this $\boldsymbol{g}$.

2.1. Let $\boldsymbol{\varphi}_1, \dots, \boldsymbol{\varphi}_n$ be a basis of S. Let M be the $n \times n$ matrix of $\mathscr{C}^1$ functions on I given by

$$M := \begin{bmatrix} \boldsymbol{\varphi}_1 & \boldsymbol{\varphi}_2 & \dots & \boldsymbol{\varphi}_n \end{bmatrix}.$$

Then, by Problem 3 of HW3, $M(t) \in GL_n(\mathbf{R})$ for $t \in I$. Thus we have a $\mathscr{C}^1$ map:

$$M \colon I \to GL_n(R).$$

Suppose $\boldsymbol{\psi}$ is a solution to (1). Setting $\boldsymbol{u} = M^{-1}\boldsymbol{\psi}$, say $\boldsymbol{u} = (u_1, \dots, u_n)$, we see that

$$\boldsymbol{\psi} = u_1\boldsymbol{\varphi}_1 + \cdots + u_n\boldsymbol{\varphi}_n.$$

We point out that $\boldsymbol{u} = (u_1, \dots, u_n)$ is not a constant vector but a map from I to

$\mathbf{R}^n$, (for $\boldsymbol{\psi}$ is not a constant vector). This means $\boldsymbol{\psi}$ need not belong to S. The map $\boldsymbol{u}\colon I \to \mathbf{R}^n$ is in $\mathscr{C}^1(I)$ as the following argument shows. The map $B \mapsto B^{-1}$ is $\mathscr{C}^\infty$ on $GL_n(\mathbf{R})$ (see related argument in (10) of § 1 of Lecture 7 of ANA2, noting that determinants and cofactors are polynomials in the coefficients and hence $\mathscr{C}^\infty$ functions of the coefficients). Since M is a $\mathscr{C}^1$ map, this means that M^{-1} is $\mathscr{C}^1$. It follows that $\boldsymbol{u} = M^{-1}\boldsymbol{\psi}$ is in $\mathscr{C}^1(I)$, since $\boldsymbol{\psi} \in \mathscr{C}^1(I)$.

Now

$$\dot{M} = \begin{bmatrix}\dot{\boldsymbol{\varphi}}_1 & \dot{\boldsymbol{\varphi}}_2 & \dots & \dot{\boldsymbol{\varphi}}_n\end{bmatrix} = \begin{bmatrix}A\boldsymbol{\varphi}_1 & A\boldsymbol{\varphi}_2 & \dots & \boldsymbol{\varphi}_n A\end{bmatrix} = AM. \tag{2.1.1}$$

Hence

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(M\boldsymbol{u}) &= \dot{M}\boldsymbol{u} + M\dot{\boldsymbol{u}} \qquad \text{(by Problem 4 of HW3)}\\ &= AM\boldsymbol{u} + M\dot{\boldsymbol{u}} \qquad \text{(by (2.1.1))}.\end{aligned} \tag{2.1.2}$$

On the other hand, since $\boldsymbol{\psi} = M\boldsymbol{u}$ is a solution of (1), we have

$$\frac{\mathrm{d}}{\mathrm{d}t}(M\boldsymbol{u}) = A(M\boldsymbol{u}) + \boldsymbol{g}.$$

Comparing this with (2.1.2) we get

$$M\dot{\boldsymbol{u}} = \boldsymbol{g}.$$

Thus $\boldsymbol{u}(t) = \int (M(t))^{-1}\boldsymbol{g}(t)dt$, where the integral represents any primitive of $M^{-1}\boldsymbol{g}$ in I. A particular solution then is

$$\boldsymbol{\psi} = M\Phi \tag{2.1.3}$$

where $\Phi\colon I \to \mathbf{R}^n$ is any primitive (i.e. anti-derivative) of $M^{-1}\boldsymbol{g}$ on I.

Remark 2.1.4. Here is a check for our calculations. Let Φ be any primitive of $M^{-1}\boldsymbol{g}$ on I. Then setting $\boldsymbol{\psi} = M\Phi$ we see that

$$\dot{\boldsymbol{\psi}} = \dot{M}\Phi + M\dot{\Phi} = AM\Phi + M(M^{-1}\boldsymbol{g}) = A\boldsymbol{\psi} + \boldsymbol{g}$$

showing that $\boldsymbol{\psi}$ is a solution of (1).

If Φ^* is another primitive of $M^{-1}\boldsymbol{g}$ on I, then $\boldsymbol{\psi}^* = M\Phi^*$ is also a solution of $\dot{\boldsymbol{x}} = A\boldsymbol{x} + \boldsymbol{g}$ from the above considerations. Therefore, from (1.1.4), $\boldsymbol{\psi}$ and $\boldsymbol{\psi}^*$ differ by an element of S. There is another way of seeing this. Since Φ and Φ^* are both primitives of $M^{-1}\boldsymbol{g}$, there exists a unique constant vector $\boldsymbol{c} = (c_1, \dots, c_n) \in \mathbf{R}^n$ such that $\Phi^* = \Phi + \boldsymbol{c}$. Now $\boldsymbol{\psi}^* = M\Phi^* = M(\Phi + \boldsymbol{c}) = M\boldsymbol{c} + M\Phi = M\boldsymbol{c} + \boldsymbol{\psi}$. Now $M\boldsymbol{c}$ is a solution of the homogeneous DE associated to (1), namley the DE $\dot{x} = A\boldsymbol{x}$, and therefore $M\boldsymbol{c} \in S$. Thus $\boldsymbol{\psi}^* - \boldsymbol{\psi} \in S$.

References

[A1] V. I. Arnold, *Ordinary Differential Equations*, translated by Richard A. Silverman, MIT Press (also Prentice-Hall, India), Cambridge,MA, U.S.A., 1973.

[A2] V. I. Arnold, *Ordinary Differential Equations*, Third Edition, translated by Roger Cooke, Universitext, Springer-Verlag, Berlin, 2006.