11. Linear Differential Equations

Consider a differential equation of the form

$$\dot{x} = A(t)x + h(t) \tag{11.1}$$

where $A(t)$ is a continuous real or complex $n \times n$ matrix valued function and $h(t)$ is a continuous n vector valued function. We assume that $A(t), h(t)$ are defined for all $t \in (-\infty, \infty)$. Let $\mathbb{V} = \mathbb{R}^n$ or $\mathbb{C}^n$. We assume both $A(\cdot), h(t)$ are $\mathbb{V}$-valued.

Given an equation (11.1), the associated *homogeneous equation* is the equation

$$\dot{x} = A(t)x \tag{11.2}$$

It can be proved that every initial value problem for (11.1) has solutions defined for all t.

Fact 11.1.
(1) *If $x(t), y(t)$ are solutions of (11.1) with $x(t_0) = y(t_0)$, then $z(t) = y(t) - x(t)$ is a solution of (11.2) with $z(t_0) = 0$.*
(2) *If $x(t)$ is a solution of (11.1) with $x(t_0) = x_0$, and $z(t)$ is a solution of (11.2) with $z(t_0) = 0$, then $y(t) = x(t) + z(t)$ is a solution of (11.2) with $y(t_0) = x_0$.*

Thus, if one knows all solutions of (11.2), and a particular solution of (11.1), then one can get all solutions of (11.1).

The *general solution* of (11.1) is a vector valued expression

$$\phi(t, c) \tag{11.3}$$

involving a vector $c = (c_1, \ldots, c_n)$ of constants so that every solution can be represented as (11.3) for a unique choice of the vector c.

We now study the general properties of equation (11.2).

Proposition 11.2. *The set of solutions to (11.2) form an n-dimensional linear subspace of the vector space of C^1 dunctions from $\mathbb{R}$ to $\mathbb{V}$.*

Proof. Let $\mathcal{S}$ be the set of solutions. Clearly, $\mathcal{S} \subset C^1(\mathbb{R}, \mathbb{V})$.

(a) $\mathcal{S}$ is a linear subspace: Suppose $x(t), y(t)$ are in $\mathcal{S}$, and a, b are scalars.

Then, $z(t) = ax(t) + by(t)$ satisfies

$$\dot{z}(t) = a\dot{x}(t) + b\dot{y}(t) = aA(t)x(t) + bA(t)y(t) = A(t)z(t),$$

so, $z(t) \in \mathcal{S}$.

(b) $\mathcal{S}$ is n-dimensional: We need to find n solutions $x_1(t), \dots, x_n(t)$ such that every solutions can be uniquely expressed as

$$x(t) = \sum_{i=1}^{n} \alpha_i x_i(t),$$

where α_i are scalars.

Let $x_i(t)$ be the unique solution such that $x_i(0) = e_i$ where e_i is the i-th standard basis vector of $\mathbb{V}$.

Let $x(t)$ be an arbitrary solution of (11.2). Then, there are scalars α_i such that $x(0) = \sum_i \alpha_i e_i$. Consider $y(t) = \sum_i \alpha_i x_i(t)$.

Then, both $x(\cdot), y(\cdot)$ are solutions and they agree at $t = 0$. By uniqueness of solutions, we have $x(t) = y(t)$ for all t.

Now, if $x(t) = \sum_i \alpha_i x_i(t) = \sum_i \beta_i x_i(t)$ for all t, then this is true for $t = 0$, so $\alpha_i = \beta_i$ $\forall i$ since $\{e_i\}$ is a basis for $\mathbb{V}$. □

A set $\{y_1(t), \dots, y_\ell(t)\}$ of solutions to (11.2) is called *linearly independent* if it is a linearly independent subset of $C^1(\mathbb{R}, \mathbb{V})$. That is, whenever

$$\sum_{i=1}^{\ell} \alpha_i y_i(t) = 0 \qquad \forall t,$$

we have $\alpha_i = 0$ $\forall i$.

Since the set of solutions to (11.2) form an n-dimensional linear subspace of the vector space of $C^1(\mathbb{R}, \mathbb{V})$, there are at most n linearly independent solutions.

Definition 11.1. *A set of n solutions $\{y_1, \dots, y_n\}$ to (11.2) is called a* fundamental set *of solutions if it is a linearly independent set. The matrix $\Phi(t) = (y_1, \dots, y_n)$, whose columns are the y_i's, is called a* fundamental matrix *for the equation (11.2) if $\{y_1, \dots, y_n\}$ is a fundamental set of solutions.*

Thus, a set $\{y_1(t), \dots, y_n(t)\}$ is a fundamental set of solutions if and only if the set form a basis for the subspace $\mathcal{S}$ of $C^1(\mathbb{R}, \mathbb{V})$.

We now want a criterion for a set $\{y_1(t), \dots, y_n(t)\}$ of solutions to be a fundamental set.

For a set $\{v_1, \dots, v_n\}$ of n vectors in a linear space $\mathbb{V}$ to be linearly independent it is necessary and sufficient that

$$\det(v_1, \dots, v_n) \neq 0$$

But linear independence of functions is a slightly different condition. The values of a set of vector valued functions $\{z_1(t), \dots, z_n(t)\}$ might be a linearly independent set of vectors for some $t's$ and be a linearly dependent set for other $t's$. Thus, if we form the function

$$W(t) = W(z_1(t), \dots, z_n(t)) = \det(z_1(t), \dots, z_n(t)),$$

we may have $W(t) = 0$ for some $t's$ and not zero for other $t's$. It turns out that this cannot happen if the $z_i(t)$ are all solutions of the same homogeneous linear differential equation (11.2).

Definition 11.2. *The determinant* $W(t) = W(y_1(t), \dots, y_n(t))$ *of the set of solutions* $\{y_1, \dots, y_n\}$ *is called the* Wronskian *of this set of solutions.*

Recall that the *trace* of an $n \times n$ matrix $A = (a_{ij})$ is

$$\text{tr}\, A = a_{11} + a_{22} + \cdots + a_{nn} = \sum_{i=1}^{n} a_{ii}. \tag{11.4}$$

Proposition 11.3. *Suppose* $\{y_1(t), \dots, y_n(t)\}$ *are* n *solutions to (11.2), and let* $W(t) = \det(y_1(t), \dots, y_n(t))$*. Then, for any real number* t_0*,*

$$W(t) = W(t_0) \exp\left(\int_{t_0}^{t} \text{tr}\, A(s) ds \right),$$

where $\text{tr}\, A(t)$ *denotes the trace of the matrix* $A(t)$*.*

Corollary 11.4. *Under the hypotheses of the proposition, if* $W(t) = 0$ *at a single* $t = t_0$*, then* $W(t)$ *is identically equal to zero.*

We will give two diferent proofs. To proceed to the proofs of the proposition, we need to use some facts about determinants, which are stated in Appendix in the end of this section.

For a matrix valued funtion $A(t) = (a_{ij}(x))$, we denote by $A'_k(t)$ the matrix valued funtion obtaained from $A(t)$ with the kth row differentiated.

Proof 1 of Propostion 11.3. Let $\Phi(t) = (y_1(t), \dots, y_n(t))$ be the matrix whose column vectors are $y_1(t), \dots, y_n(t)$. Note that if $y_i(t) =$

$\begin{pmatrix} y_{1i}(t) \\ \dots \\ y_{ni}(t) \end{pmatrix}$ is a solution of (11.2), then $\dot{y}_i(t) = \begin{pmatrix} \sum_{\ell=1}^n a_{1\ell}(t) y_{\ell i}(t) \\ \dots\dots \\ \sum_{\ell=1}^n a_{n\ell}(t) y_{\ell i}(t) \end{pmatrix}$. So

$$\det \Phi_1'(t) = \det \begin{pmatrix} \dot{y}_{11}(t) & \dots & \dot{y}_{1n}(t) \\ y_{21}(t) & \dots & y_{2n}(t) \\ \dots\dots & \dots & \dots\dots \\ y_{n1}(t) & \dots & y_{nn}(t) \end{pmatrix}$$

$$= \det \begin{pmatrix} \sum_{\ell=1}^n a_{1\ell}(t) y_{\ell 1}(t) & \dots & \sum_{\ell=1}^n a_{1\ell}(t) y_{\ell n}(t) \\ y_{21}(t) & \dots & y_{2n}(t) \\ \dots\dots & & \\ y_{n1}(t) & \dots & y_{nn}(t) \end{pmatrix} = a_{11}(t) W(t),$$

where we use the properties of determinants stated in Lemma 11.7.

Similarly we can get the equality for $\det \Phi_i'(t)$, $i = 2, \dots, n$,

$$\det \Phi_i'(t) = a_{ii}(t) W(t).$$

Hence by Lemma 11.8 in Appendix, we can get

$$\frac{d}{dt} \det W(t) = \big(\sum_{i=1}^n a_{ii}(t)\big) W(t) = \mathrm{tr}(A(t)) W(t).$$

Now it follows that $W(t) = W(t_0) \exp \left\{ \int_{t_0}^t \mathrm{tr}\, A(s) ds \right\}$. □

Proof 2 of Propostion 11.3. Let $\Phi(t)$ be the matrix whose columns are the solutions $y_1(t), \dots, y_n(t)$, so that we have the matrix equation

$$\Phi'(t) = A(t)\Phi(t).$$

Now, we have

$$\begin{aligned} \Phi(t+h) &= \Phi(t) + \Phi'(t)h + r(t,h) \\ &= \Phi(t) + A(t)\Phi(t)h + r(t,h) = (I + A(t)h)\Phi(t) + r(t,h), \end{aligned}$$

where $r(t,h)$ is a matrix valued funtion satisfying

$$\lim_{h \to 0} \frac{1}{h} r(t,h) = 0. \tag{11.5}$$

Hence, we have

$$W(t+h) = \det\big((I + A(t)h)\Phi(t) + r(t,h)\big).$$

Applying Corollary 11.11 for the matrices $(I + A(t)h)\Phi(t)$ and $r(t,h)$, we get

$$\det\big((I + A(t)h)\Phi(t) + r(t,h)\big) = \det\big((I + A(t)h)\Phi(t)\big) + r_1(t,h),$$

where $r_1(t,h)$ is a real function satisfying $|r_1(t,h)| \leq K|r(t,h)|$. Hence we have $\lim_{h\to 0} r_1(t,h)/h = 0$ because of (11.5).

Using the fact that $\det(AB) = \det(A)\det(B)$ for all matrices A, B, Lemma 11.9 gives

$$W(t+h) = W(t)(1 + h\,\mathrm{tr}\,A(t) + O(h^2)) + r_1(t,h)$$

So if $W(t) \neq 0$, then

$$\frac{W(t+h) - W(t)}{h} = W(t)\,\mathrm{tr}\,A(t) + \frac{O(h^2)}{h} + \frac{r_1(t,h)}{hW(t)},$$

which implies $W'(t) = (\mathrm{tr}\,A(t))W(t)$. If $W(t) = 0$, then we have $W(t+h) = r_1(t,h)$, and

$$\frac{W(t+h) - W(t)}{h} = \frac{r_1(t,h)}{h},$$

which also gives $W'(t) = 0 = (\mathrm{tr}\,A(t))W(t)$. This is what we need. □

The above proposition gives that if the Wronskian is not zero for some t, then it is not zero for any t.

Since the solution set of equation (11.2) is of n dimensional, a fundamental set of solutions determines all solutions of (11.2). In fact, the general solution to (11.2) has the form

$$x(t) = \Phi(t)c,$$

where $\Phi(t)$ is any fundamnental matrix for (11.2) and c is a constant vector.

By the following theorem, the general solution of (11.1) can also be obtained form a fundamental set of solutions of the corresponding homogeneous equation (11.2).

Theorem 11.5. *Suppose $\Phi(t)$ is a fundamental matrix for (11.2). Then any particular solution $x(t)$ for (11.1) has the form*

$$x(t) = \Phi(t)c + \Phi(t)\int_{t_0}^{t} \Phi(s)^{-1}h(s)ds,$$

where $c \in \mathbb{R}^n$ is a vector satisfying $x(t_0) = \Phi(t_0)c$.

Proof. We look for solution of the form

$$x_p(t) = \Phi(t)v(t),$$

where $v(t)$ is some non-constant vector-valued function of t.

Then,

$$x_p' = \Phi' v + \Phi v' = A\Phi v + \Phi v'.$$

On the other hand,

$$x_p' = Ax_p + h = A\Phi v + h.$$

So we should have

$$\Phi v' = h.$$

Since Φ is invertible, we can write the equation as

$$v' = \Phi(t)^{-1}h(t).$$

Integrating, gives

$$v(t) = c + \int_{t_0}^{t} \Phi(s)^{-1}h(s)ds,$$

where c is chosen such that $x(t_0) = \Phi(t_0)c$.

Now, take $v(t)$ in this equation and reverse the steps. This gives us

$$x(t) = \Phi(t)c + \Phi(t)\int_{t_0}^{t} \Phi(s)^{-1}h(s)ds$$

□

This last formula is known as the *variation of constants formula* or the *variation of parameters* formula.

Appendix

For a matrix valued funtion $A(t) = (a_{ij}(x))$, we denote by $A_k'(t)$ the matrix valued funtion obtaained from $A(t)$ with the kth row differentiated.

Let $A = (a_{ij})$ be an $n \times n$ matrix. Let M_{ij} denote the $(n-1)\times(n-1)$ matrix obtained by deleting the ith row and jth column of A. The *minor* of a_{ij} is the determinant of M_{ij}, $\det M_{ij}$, and the *cofactor* of a_{ij} is $A_{ij} = (-1)^{i+j}\det M_{ij}$.

The *Laplace expansion*, or *cofactor expansion* of the determinant of a matrix is given by the following theorem.

Theorem 11.6. *Suppose $A = (aij)$ is an $n \times n$ matrix. Then its determinant $\det A$ is given by*

$$\begin{aligned}\det A =& a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}\\ =& a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj}\\ =& \sum_{k=1}^{n} a_{ik}A_{ik} = \sum_{k=1}^{n} a_{kj}A_{kj}.\end{aligned}$$

The following properties for determinants are well known.

Lemma 11.7. *Let A be a matrix.*

(1) *If a matrix B is obtained from A by multiplying a single row or column of A by a constant c, then*
$$\det(B) = c\det(A).$$
(2) *If a matrix B is obtained from A by adding a multiple of one row or column to another row or column, respectively, then*
$$\det(B) = \det(A).$$

For a matrix valued funtion $A(t) = (a_{ij}(t))$, we denote by $A'_k(t)$ the matrix valued funtion obtaained from $A(t)$ with the kth row differentiated.

Lemma 11.8. *Let $A(t) = (a_{ij}(t))$ be an $n \times n$ matrix valued funtion. Then*
$$\frac{d}{dt}\det A(t) = \det A'_1(t) + \det A'_2(t) + \cdots + \det A'_n(t) = \sum_{k=1}^{n} \det A'_k(t).$$

Proof. If $n = 2$ and $A(t) = \begin{pmatrix} a_{11}(t) & a_{12}(t) \\ a_{21}(t) & a_{22}(t) \end{pmatrix}$. Then
$$\begin{aligned}
&\frac{d}{dt}\det A(t) = \frac{d}{dt}(a_{11}(t)a_{22}(t) - a_{12}(t)a_{21}(t)) \\
=&(\dot{a}_{11}(t)a_{22}(t) - \dot{a}_{12}(t)a_{21}(t)) + (a_{11}(t)\dot{a}_{22}(t) - a_{12}(t)\dot{a}_{21}(t)) \\
=&\begin{vmatrix} \dot{a}_{11}(t) & \dot{a}_{12}(t) \\ a_{21}(t) & a_{22}(t) \end{vmatrix} + \begin{vmatrix} a_{11}(t) & a_{12}(t) \\ \dot{a}_{21}(t) & \dot{a}_{22}(t) \end{vmatrix} = \det A'_1(t) + \det A'_2(t)
\end{aligned}$$

For $n > 2$, we can use Laplace expansion for determinant
$$\det A(t) = a_{11}(t)A_{11}(t) + \cdots + a_{1n}(t)A_{1n}(t),$$
where $A_{1j}(t)$ is the cofactor of $a_{1j}(t)$, and then use induction. □

Recall that the trace of an $n \times n$ matrix $A = (a_{ij})$ is given in 11.4.

Let I_n denote the $n \times n$ identity matrix.

Let $O(x)$ denote any function $R(x)$ such that there is a constant $C > 0$ such that
$$\limsup_{x\to 0} \frac{R(x)}{x} < C.$$

Lemma 11.9. *Let $A = (a_{ij})$ by an $n \times n$ matrix, and let $h > 0$. Then,*
$$\det(I_n + hA) = 1 + h\operatorname{tr} A + O(h^2).$$

Proof. We use induction of n.

It is trivial for $n = 1$ since we can use the zero function for $O(h^2)$.

Assume the Lemma is true for $n - 1$.

Let $B = I_n + Ah$.

Using cofactor expansion down the first column of B gives

$$\det(I_n + Ah) = (1 + a_{11}(t)h)B_{11} + \sum_{i=2}^{n}(-1)^{1+j}ha_{i1}B_{i1},$$

where B_{ij} is the cofactor of b_{ij} for matrix B, and has the form $B_{i1} = (-1)^{i+1}M_{i1}(B)$, where $M_{i1}(B)$ is obtained by deleting the ith row and the first column of B.

It is clear that

$$M_{11}(B) = I_{n-1} + hM_{11}(A).$$

So, by induction, we have the first entry on the right side of () equals

$$\begin{aligned}&(1 + ha_{11}(t))(1 + h\,\mathrm{tr}\,M_{11}(A) + O(h^2))\\ =&1 + ha_{11}(t) + h\,\mathrm{tr}\,M_{11}(A) + O(h^2)\\ =&1 + h\ \ \mathrm{tr}\,A + O(h^2).\end{aligned}$$

On the other hand, each entry in the sum in the second term in equation () has an h in its first column. So, this whole sum is $O(h^2)$. ☐

We need a standard result from the calculus of maps from $\mathbb{R}^n$ to $\mathbb{R}$.

Lemma 11.10. *Let $\psi : D \to \mathbb{R}$ be a C^1 function defined on an open set $D \subset \mathbb{R}^n$, and let u be vectors in D. Then for any bounded neighborhood U of u such that $\overline{U} \subset D$, there is a constant $K > 0$ such that for any $v \in D$ with $\{u + tv : 0 \le t \le 1\} \subset U$,*

$$\mid \psi(u + v) - \psi(u) \mid \le K \mid v \mid .$$

Proof. Let

$$K = \max\{\|D\psi_w\| : w \in U\}.$$

By continuity of $D\psi$ and compactness of $\overline{U}$, $K < \infty$. Now we have

$$\psi(u + v) - \psi(u) = \int_0^1 \frac{d}{dt}\psi(u + tv)dt = \int_0^1 D\psi_{u+tv} \cdot vdt,$$

and therefore

$$\|\psi(u + v) - \psi(u)\| \le \int_0^1 \|D\psi_{u+tv}\|dt \cdot \|v\| \le K\|v\|$$

by the choice of K. ☐

We now apply this to the determinant function, $\det(\cdot)$, on $n \times n$ matrices. We can think that matrices are elements of $\mathbb{R}^{n^2}$, and define the norm of an matrix by

$$| A |=| A |_2= (\sum_{i,j} a_{ij}^2)^{\frac{1}{2}}.$$

Corollary 11.11. *Let A be an $n \times n$ matrices. Then for any $\epsilon > 0$, there exist a constant $K > 0$ such that for any matrix B with $| B |_2 < \epsilon$,*

$$| \det(A + B) - \det(A) | \leq K | B |_2$$

Proof. This is because the function $\det : \mathbb{R}^{n^2} \to \mathbb{R}^1$ is a C^1 function. Then we use the above lemma. $\square$