Linear Differential Equations

Consider a differential equation of the form

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

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 $V = \mathbf{R}^n$ or $\mathbf{C}^n$. We assume both $A(\cdot), h(t)$ are $V - valued$.

Given an equation (1), the associated homogeneous equation is the equation

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

We have already proved that every initial value problem for (1) has solutions defined for all t.

Basic Facts:

1. If $x(t), y(t)$ are solutions of (1), then $z(t) = y(t) - x(t)$ is a solution of (2).
2. If $x(t)$ is a solution of (1) and $z(t)$ is a solution of (2), then $y(t) = x(t) + z(t)$ is a solution of (1).

Thus, if one knows all solutions of (2), and one knows one particular solution of (1), then one can get all solutions of (1).

The *general solution* of (1) is a vector-valued expression

$$\phi(t, \mathbf{c}) \tag{3}$$

involving a vector $\mathbf{c} = (c_1, \ldots, c_n)$ of constants so that every solution can be represented as (3) for a unique choice of the vector $\mathbf{c}$.

We now study the general properties of equation (2).

Proposition. *The set of solutions to (2) form an $n-$dimensional linear subspace of the vector space of C^1 dunctions from $\mathbf{R}$ to V.*

Proof.

Let $\mathcal{S}$ be the set of solutions.

Clearly, $\mathcal{S} \subset C^1(\mathbf{R}, V)$.

(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

$$\begin{aligned} \dot{z}(t) &= a\dot{x}(t) + b\dot{y}(t) \\ &= aA(t)x(t) + bA(t)y(t) \\ &= A(t)(z(t)) \end{aligned}$$

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

(B) $\mathcal{S}$ is $n-$dimensional:

We need to find n solutions $x_1(t), \ldots, x_n(t)$ such that every solutions can be uniquely expressed as

$$x(t) = \sum_{i=1}^{n} \alpha_i x_i(t) \tag{4}$$

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 V.

Let $x(t)$ be an arbitrary solution of (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 V. QED

Definition. A set $\{y_1(t), \ldots, y_\ell\}$ of solutions to (2) is called linearly independent if it is a linearly independent subset of $C^1(\mathbf{R}, V)$. That is, whenever

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

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

A maximal linearly independent set of solutions is called *a fundamental set* of solutions.

Thus, a set $\{y_1(t), \ldots, y_n(t)\}$ is a fundamental set of solutions iff the set is a basis for the subspace $\mathcal{S}$ of $C^1(\mathbf{R}, V)$.

Note, that, since any two bases of a vector space have the same number of elements, we have $\ell = n$ for any fundamental set of solutions.

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

For a set $\{v_1, \ldots, v_n\}$ of vectors in V to be linearly independent if is necessary and sufficient that

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

But linear independence of functions is a slightly different condition. The values of a set of vector valued functions $\{z_1(t), \ldots, z_n(t)\}$ 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

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

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 (2).

Proposition. *Suppose* $\{y_1(t), \ldots, y_n(t)\}$ *are* n *solutions to (2), and let* $W(t) = det(y_1(t), \ldots, y_n(t))$.

Then, for any real t_0,

$$W(t) = W(t_0)exp\left(\int_{t_0}^{t} trace(A(s))ds\right) \tag{5}$$

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

Before we proceed to the proof of the proposition, we recall some properties of determinants.

If $A = (a_{ij})$ is an $n \times n$ matrix, let $A(i \mid j)$ denote the $(n-1) \times (n-1)$ matrix obtained by deleting the $i-$th row and $j-$th column of A.

The expanding by minors along the first column gives

$$det(A) = \sum_{j=1}^{n} (-1)^{1+j} a_{j1} det A(j \mid 1). \tag{6}$$

Also, if h is a real number, and B is obtained from A by multiplying a single row (or column) of A by h, then

$$det(B) = h\ det(A). \tag{7}$$

We define $tr(A) = trace(A) = \sum_{i=1}^{n} a_{ii}$ to be the sum of the diagonal elements of A.

Lemma 1. Let $A(t) = (a_{ij}(t))$ by a continuous $n \times n$ matrix function of t, and let $h > 0$. Then,

$$det(I + A(t)h) = 1 + tr(A(t))h + O(h^2). \quad (8)$$

Here I is the $n \times n$ identity matrix and $O(h^2)$ denotes any function $R(h)$ such that there is a constant $C > 0$ such that

$$\limsup_{h \to 0} \frac{R(h)}{h^2} < C.$$

Proof of Lemma 1.
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 $B = I + A(t)h$.
Expanding by minors down the first column of B gives

$$\begin{aligned} det(I + A(t)h) &= (1 + a_{11}(t)h)det(B(1 \mid 1) \\ &+ \sum_{j=2}^{n}(-1)^{1+j} h a_{j1}(t) det(B(j \mid 1). \end{aligned} \quad (9)$$

Now, letting I_{n-1} denote the $(n-1) \times (n-1)$ identity matrix, we have

$$B(1 \mid 1) = I_{n-1} + hA(1 \mid 1).$$

So, by induction, we have the first entry on the right side of (9) equals

$$\begin{aligned} (1 + a_{11}(t)h)(1 + h\ trA(1 \mid 1) + O(h^2)) &= 1 + a_{11}(t)h + \\ &\quad h\ trA(1 \mid 1) + O(h^2) \\ &= 1 + h\ trA(t) + O(h^2). \end{aligned}$$

On the other hand, each entry in the sum in the second term in equation (9) has an h in its first column. So, this whole sum is $O(h^2)$. QED.

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

Lemma 2. Let $\psi : \mathbf{R}^n \to \mathbf{R}$ be a C^1 function defined on an open set $D \subset \mathbf{R}^n$, and let u, v be vectors in D such that the closed line segment joining

u to $u+v$ is completely contained in D (i.e. $\{u+tv : 0 \leq t \leq 1\} \subset D$). Then, there is a constant $K > 0$ such that

$$| \psi(u+v) - \psi(u) | \leq K | v |. \tag{10}$$

Proof.

Let $\nabla\psi(x)$ denote the gradient of ψ at the point x.

Since the map $t \to | \nabla\psi(u+tv) |$ is continuous on the closed unit interval $[0,1]$, we have that

$$K \stackrel{\text{def}}{=} \sup_{0\leq t\leq 1} | \nabla\psi(u+tv) | < \infty.$$

The map $\eta(t) = \psi(u+tv)$ is a C^1 real-valued function of the variable t on the closed interval $[0,1]$, so the Mean-Value theorem gives that there is a $0 < t_0 < 1$ such that

$$\psi(u+v) - \psi(u) = \eta(1) - \eta(0) = \eta'(t_0)(1-0) = \eta'(t_0).$$

By the Chain Rule for functions of several variables, we have that

$$\eta'(t_0) = \nabla\psi(u+t_0v) \cdot (v)$$

where $\cdot$ denotes the usual dot product (inner product) in $\mathbf{R}^n$.

By the Cauchy-Schwarz inequality, we have

$$\begin{aligned} | \psi(u+v) - \psi(u) | &= | \eta'(t_0) | \\ &= | \nabla\psi(u+t_0v) \cdot (v) | \\ &\leq | \nabla\psi(u+t_0v) || \cdot(v) | \\ &\leq K| v | \end{aligned}$$

QED.

We now apply this to the determinant function, *det*, on $n \times n$ matrices, thinking of these as elements of $\mathbf{R}^{n^2}$.

Lemma 3. Let A and B be two $n \times n$ matrices. Then, there is a constant K such that

$$| det(A+B) - det(A) | \leq K | B | \tag{11}$$

Note that we may write this using the 'big oh' notation as

$$det(A+B) = det(A) + O(B)$$

Proof of Lemma 3.
We recall that, if $A = (a_{ij})$, then

$$det(A) = \sum_{\text{permutations } \sigma} (-1)^{sgn(\sigma)} a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)}$$

where $sgn(\sigma)$ denotes the sign of the permutation σ.

Thus, $det(A)$ is a polynomial in the coefficients (a_{ij}). (This can also be seen using induction on n and the expansion by minors along the first column formula).

Thus, as a function on $\mathbf{R}^{n^2}$, the map $A \to det(A)$ is certainly C^1.

We apply Lemma 2. Note, that, since we used the Cauchy-Schwarz inequality, the constant K obtained in Lemma 2 requires that we use the norm

$$| A |_1 = (\sum_{i,j} a_{ij}^2)^{\frac{1}{2}}$$

on matrices A.

Thus, we get some constant, say K_1 such that

$$| det(A+B) - det(A) | \le K_1 | B |_1$$

Since, any two norms on $\mathbf{R}^{n^2}$ are boundedly related, we have that,

$$| B |_1 \le K_2 | B |$$

where $| B |$ is the usual operator norm $\sup_{| x |=1} | Bx |$.

Thus, (11) follows using $K = J_1 K_2$. QED.

Proof of the proposition.
We show that $W(t)$ satisfies the scalar linear non-autonomous differential equation

$$W' = trA(t)W \tag{12}$$

It will then follow that

$$\frac{W'}{W} = trA(t)$$

or

$$\log W(t) \mid_{W(t_0)}^{W(t)} = \int_{t_0}^{t} trA(s)ds$$

or

$$W(t) = W(t_0)exp\left(\int_{t_0}^{t} trA(s)ds\right)$$

As usual, we let $o(h)$ denote a function $r(h)$ such that

$$\lim_{h\to 0} \frac{r(h)}{h} = 0.$$

Let $\Phi(t)$ be the matrix whose columns are the solutions $y_1(t), \ldots, y_n(t)$, so that we have the matrix equation

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

Now, we have

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

which, in turn, gives

$$W(t+h) = det((I + A(t)h)\Phi(t) + o(h))$$

Applying Lemma 3 with $A = (I + A(t)h)\Phi(t)$ and $B = o(h)$, we get

$$\begin{aligned} det((I + A(t)h)\Phi(t) + o(h)) &= det((I + A(t)h)\Phi(t)) + O(o(h)) \\ &= det((I + A(t)h)\Phi(t)) + o(h) \end{aligned}$$

Using the fact that the $det(AB) = det(A)det(B)$ for all A, B and Lemma 1 gives

$$\begin{aligned} W(t+h) &= W(t)(1 + tr\ A(t)h + O(h^2)) + o(h) \\ &= W(t) + W(t)(tr\ A(t)h) + o(h). \end{aligned}$$

So,

$$\frac{W(t+h) - W(t)}{h} = W(t)tr\ A(t) + \frac{o(h)}{h}.$$

Letting $h \to 0$ gives

$$W' = trA(t)W$$

as required. QED.

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

This a set of solutions is a fundamental set if and only if its Wronskian is not zero for some (or any) t.

When, $\{y_1, \ldots, y_n\}$ is a fundamental set of solutions, we form the matrix $\Phi(t)$ whose columns are the $y_i's$ and call this matrix a *fundamental matrix* for (2).

Thus, the general solution to (2) has the form

$$x(t) = \Phi(t)\mathbf{c}$$

where $\Phi(t)$ is any fundamnental matrix for (2) and $\mathbf{c}$ is a constant vector.

Fact: Suppose $\Phi(t)$ is a fundamental matrix for (2) and $h(t)$ is not identically zero. Then, one can always find a particular solution for (1) of the form

$$x_p(t) = \Phi(t)v(t) \tag{13}$$

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

Proof.

To motivate how we find $v(t)$, first suppose that we can find a solution as in (13).

Then,

$$\begin{aligned} x_p' &= \Phi' v + \Phi v' \\ &= A\Phi v + \Phi v' \\ &= Ax_p + h \\ &= A\Phi v + h \end{aligned}$$

Or,

$$\Phi v' = h$$

Since, Φ is invertible, we can write the last equation as

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

Integrating, gives

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

for any t_0.
Now, take $v(t)$ in this last equation and reverse the steps. This gives us

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

and, the general solution to (1) has the form

$$x(t) = \Phi(t)\mathbf{c} + \Phi(t) \int_{t_0}^{t} \Phi(s)^{-1} h(s) ds \tag{14}$$

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